

Remember/Know Judgments Probe Degrees of Recollection

Peter E. Wais, Laura Mickes, and John T. Wixted

Abstract

■ Remembering and knowing are states of awareness that accompany the retrieval of facts, faces, and experiences from our past. Although originally intended to separate episodic from semantic memory, the dominant view today is that recollection-based decisions underlie remember responses, whereas familiarity-based decisions underlie know responses. Many functional magnetic resonance imaging (fMRI) studies as well as lesion studies have relied on the remember/know procedure to identify the neural correlates of recollection and familiarity. An implicit assumption of this approach is that know responses, which are thought to tap familiarity-based decisions, are devoid of recollection. We investigated this issue

by using a source memory procedure and found that the accuracy of source recollection was significantly above chance for studied words that were declared to be old and known. Critically, this held true even when the source decision was made before the old/new decision (i.e., even after successful recollection had just occurred). Our results show that although recollection and familiarity may be different processes, the remember/know paradigm does not probe them directly. As such, dissociations involving remember/know judgments in fMRI studies and in studies involving amnesic patients should not be construed as dissociations between recollection and familiarity. ■

INTRODUCTION

Dual-process theory holds that some recognition memory decisions are based on the recollection of contextual detail, whereas other decisions are based on a sense of familiarity unaccompanied by contextual information (Curran & Hintzman, 1995; Mandler, 1980). One technique that is designed to measure these two processes is the remember/know procedure, which simply involves asking participants to say “remember” (R) for recollection-based decisions and “know” (K) for familiarity-based decisions (Gardiner, Richardson-Klavehn, & Ramponi, 1997). Many recent neuroimaging studies have used this convenient procedure to investigate the neural correlates of recollection and familiarity (e.g., Otten, in press; Eldridge, Engel, Zeineh, Bookheimer, & Knowlton, 2005; Gonsalves, Kahn, Curran, Norman, & Wagner, 2005¹; Uncapher & Rugg, 2005; Woodruff, Johnson, Uncapher, & Rugg, 2005; Wheeler & Buckner, 2004; Eldridge, Knowlton, Furmanski, Bookheimer, & Engel, 2000; Henson, Rugg, Shallice, Josephs, & Dolan, 1999; Brewer, Zhao, Desmond, Glover, & Gabrieli, 1998). The remember/know (R/K) procedure is also commonly used to investigate the process-specific effects of hippocampal lesions (e.g., Aggleton et al., 2005; Holdstock, Mayes, Gong, Roberts, & Karpur, 2005; Verfaellie, Cook, & Keane, 2003; Holdstock et al.,

2002; Moscovitch & McAndrews, 2002; Yonelinas et al., 2002; Yonelinas, Kroll, Dobbins, Lazzara, & Knight, 1998).

An explicit assumption underlying the R/K procedure is that individual recognition decisions involve one process or the other (never both processes together). An alternative view is that R/K judgments denote different levels of memory strength, not different processes. According to this view, high-strength memories (associated with R responses) are high in both recollection and familiarity, on average, whereas low-strength memories (associated with K responses) are lower in both processes, on average (Wixted, 2007). Such an account is consistent with the signal-detection interpretation of R/K judgments (Dunn, 2004; Wixted & Stretch, 2004; Donaldson, 1996).

A key difference between these two conceptualizations is how they interpret K judgments. According to most accounts, K judgments reflect familiarity-based decisions that are devoid of recollection. According to the alternative signal-detection account, they instead reflect decisions on the basis of low-strength memories that include lesser degrees of recollection (relative to R responses). One way to differentiate between these two accounts is to use a source recollection procedure. In this procedure, items are presented from one of two sources (e.g., words in either blue or red font). On the recognition test, participants are first asked to make an old/new decision (along with an R/K judgment for each item that is declared to be old) and are then asked to recollect the item’s source. If K responses reflect

University of California, San Diego

familiarity-based decisions that are devoid of recollection, then the accuracy of the subsequent source recollection decision ought to be at chance. If they instead reflect decisions that are partially based on recollection, then the accuracy of the subsequent source decision ought to be above chance (i.e., recollection accuracy should fall between that associated with R responses and chance performance).

Perfect, Mayes, Downes, and van Eijk (1996) and Conway and Dewhurst (1995) conducted source memory experiments along these lines and found that source accuracy associated with K responses was generally above chance, but not always significantly so. Eldridge et al. (2005) also found that source recollection success was significantly above chance levels for K responses and was higher still for R responses. In other words, K responses involved less recollective detail than R responses but did not signal the absence of recollective detail.

One way to reconcile above-chance source recollection for K responses with the traditional familiarity-based interpretation is to assume that the initial K response associated with the old/new decision in these experiments was, indeed, based on recollection-free familiarity but that the subsequent source decision was based on recollection that became available after the old/new judgment was completed. That is, recollection might have sometimes succeeded after a second query of memory or after additional search time that occurred between the old/new question and the source question (e.g., Gronlund, Edwards, & Ohrt, 1997; Hintzman & Caulton, 1997; Johnson, Kounios, & Reeder, 1994). The purpose of the research we conducted was to investigate that possibility and, more generally, to more clearly establish the relationship between K responses and the availability of recollective information. In the critical test, participants made a source decision *first* and then made an old/new decision that was accompanied by an R/K judgment.

METHODS

Participants

The participants were 39 college undergraduates (10 men), who were recruited from the university experimental participants pool, gave their informed consent according to the university IRB protocol, and received class credit for completing our experiment. The experiment was run in two versions, with 20 participants assigned to version 1 and 19 to version 2. All participants were native English speakers and free from the effects of any reported medication at the time they completed the experiment.

Stimuli

From a list of 200 English nouns developed in previous experiments (i.e., names of human body parts, American

states, foreign countries, and international cities), 70 target words were selected and divided into two equal lists such that Block A and Block B contained words that were as closely thematically matched as possible and balanced for word categories. The 130 remaining words were used as lures during the test session. All stimuli were presented using E-Prime 1.1.4.1 (Psychology Software Tools) scripts on a Dell Dimension 4550 desktop computer and 17-in. liquid crystal display throughout both the encoding and test conditions.

Procedure

Participants were informed that lists of words would be presented in front of them and they were instructed about how to perform their tasks for each session. Participants ran brief practice scripts before each the encoding and test sessions to ensure their familiarity with their tasks. In the encoding session, each target word was presented for 5.0 sec while the participant rated each word as pleasant or unpleasant on the computer keyboard. Words in Block A were presented above the center of the screen in blue Helvetica font, and words in Block B were presented below the center of the screen in red Tahoma font (these were the two sources). During the practice session, each participant was advised that their memory for the words, including the presentation color, would be tested when they returned for their second session.

After a 1-hr retention interval, each participant completed a recognition test. The 200 test words were presented in four equal blocks, providing the participant a 10-sec rest break in between blocks. Test words were presented in black Courier font for 5 sec during each of two test questions. The recognition test included an old/new question in one step, for which participants could answer “remember,” “know,” “guess,” or “new” as they viewed the test word, and a source recollection question in another step, for which they either entered their confidence rating about the color that the word was presented in during the encoding session, or they entered “new.” Their source confidence rating was entered on a 7-point scale (e.g., definitely, probably, maybe blue; new, definitely, probably, or maybe red). A guess (G) option was provided during the old/new phase because prior research by R/K theorists has been interpreted to mean that subjects tend to include guesses in their K responses rather than in their R responses, thereby selectively contaminating the former (e.g., Gardiner et al., 1997; Gardiner, Java, & Richardson-Klavehn, 1996).

The experiment was run in two versions. The versions were identical in terms of the study phase, but differed in regard to the test phase. In the test phase of version 1, each item was presented for an old/new decision followed by the source recollection question. In version 2, the question order was reversed so that participants

answered the source recollection question first, before answering the old/new question. Before practicing the test session, each participant received specific instructions about the use of the R, K, and G responses for the old/new task and heard the descriptions developed by Gardiner and Richardson-Klavehn (2000) verbatim. Those instructions are reproduced here in the Appendix. Participants were encouraged to have these instructions repeated to them until they were completely clear about when to apply each type of recognition response. In both experiments, it was made clear to subjects that the R/K/G judgment pertained to the item's status on the list (i.e., the question pertained to the old/new decision).

RESULTS

All participants performed well on the item recognition task, but the source recollection task proved to be more difficult. Of the 39 participants tested in the two versions of the experiment, the source recollection performance of 13 participants did not exceed chance (7 participants in version 1 and 6 participants in version 2). Of those, six were slightly (and nonsignificantly) below chance, and seven were slightly (and nonsignificantly) above chance according to a binomial test. Because the data from these participants could not help to address the main question of interest (i.e., is source accuracy for K responses greater than chance but less than that for R responses?), their results were analyzed separately.

Old/New Performance

Performance on the old/new portion of the task is shown in Table 1. The results for version 1 ($n = 13$) and for version 2 ($n = 13$) show the performance of the participants who exhibited above-chance accuracy on the source recollection question. The table also shows the combined performance of the 13 participants from both versions of the experiment who did not exhibit

above-chance source recollection accuracy. The overall hit and false alarm rates (bottom line of the table) were similar for all three groups. Separate analyses of variance performed on the hit rates, false alarm rates, and d' scores from the three groups were all far from significant. Table 1 also shows the hit and false alarm rates broken down by response category (R, K, and G), and these were similar across groups as well.

Figure 1 shows the mean accuracy of old/new decisions as a function of response category (guess, know, or remember) for all 39 participants combined. Remember accuracy for a given participant is equal to the R hit rate divided by the sum of the R hit rate plus the R false alarm rate; Know accuracy is equal to the K hit rate divided by the sum of the K hit rate plus the K false alarm rate; and Guess accuracy is equal to the G hit rate divided by the sum of the G hit rate plus the G false alarm rate. The results show that subjects actually were guessing when they used the guess option, as their accuracy for those responses was not significantly different from chance. In addition, R responses were highly accurate, whereas accuracy for K responses fell between these two extremes.

Source Recollection

In version 1, the mean source recollection performance for the 13 participants who exhibited above-chance source accuracy was 68% correct. The corresponding value in version 2 was 65% correct, and these values did not differ significantly. Mean source recollection accuracy for the 13 participants who did not exhibit above-chance source recollection was 51%.

The key findings of this experiment concern source accuracy for old/new decisions that received G, K, or R responses. The mean source accuracy for words that received a G response during the old/new recognition test was 0.46 (0.09), which did not differ significantly from chance [$t(11) = -0.63, p < .55$]. For K responses, mean source accuracy was 0.60 (0.06), which was significantly above chance [$t(11) = 3.69, p < .01$]. One

Table 1. Mean Hit and False Alarm Rates for Version 1 ($n = 13$), Version 2 ($n = 13$), and Excluded Participants ($n = 13$) are Listed with the Standard Error of those Means

	<i>With Source Recollection</i>				<i>Without Source Recollection</i>	
	<i>Version 1</i>		<i>Version 2</i>		<i>Combined</i>	
	<i>Hits</i>	<i>False Alarms</i>	<i>Hits</i>	<i>False Alarms</i>	<i>Hits</i>	<i>False Alarms</i>
Remember	0.58 (0.06)	0.02 (0.01)	0.66 (0.06)	0.05 (0.01)	0.65 (0.05)	0.06 (0.01)
Know	0.24 (0.06)	0.05 (0.01)	0.15 (0.03)	0.07 (0.02)	0.14 (0.02)	0.08 (0.01)
Guess	0.06 (0.01)	0.11 (0.02)	0.09 (0.02)	0.09 (0.02)	0.07 (0.02)	0.10 (0.03)
Old/New	0.88 (0.02)	0.18 (0.03)	0.90 (0.02)	0.21 (0.03)	0.86 (0.03)	0.24 (0.04)

Hits represent the proportion of targets endorsed as old, and false alarms represent the proportion of lures endorsed as old.

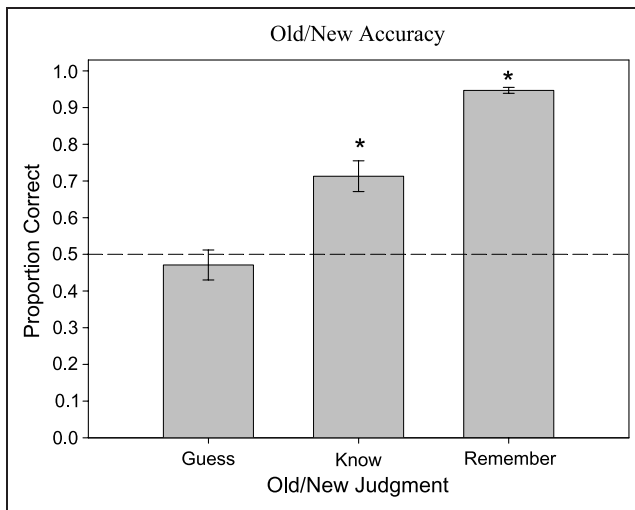


Figure 1. Old/new accuracy for remember, know, and guess responses collapsed across all subjects ($n = 39$). For both remember and know responses, mean proportions with correct source judgments were significantly above the 50% chance level ($p < .01$).

participant who did not make any K responses (and therefore who had no source accuracy score for K responses) was not included in this analysis. For R responses, mean source accuracy was 0.72 (0.03), which was also significantly above chance [$t(12) = 7.45, p < .01$].

Similar results were observed in version 2. For G responses, source accuracy was 0.51 (0.08), a value that was not significantly different from chance. For K responses, source accuracy was 0.59 (0.06), which was significantly above chance [$t(11) = 3.70, p < .01$]. Once again, one participant who did not make any know responses was not included in this analysis. For R responses, source accuracy was 0.68 (0.02), which was also significantly above chance [$t(12) = 7.53, p < .01$].

The two versions of the experiment did not produce significantly different source accuracy results for G, K, or R responses, so the data from the two versions were collapsed for further analysis. Figure 2 shows source accuracy as a function of the judgment associated with the old/new decision (guess, know, and remember) for the 26 participants who exhibited above-chance source recollection. This figure summarizes the main result of our study. The critical finding is that these data show that recollective success increases monotonically across the three judgments. A linear trend analysis as a function of memory strength (with guess, know, and remember representing low, medium, and high strength) was highly significant [$F(1, 21) = 12.86, p < .01$].

Not surprisingly, no such effects were evident for the 13 participants whose overall source accuracy did not exceed chance. Although their old/new recognition memory performance was similar to that of the other subjects (as shown in Table 1), they clearly did not encode the source information. For these participants, source accuracy for G, K, and R responses was 0.42, 0.49,

and 0.52, respectively, none of which differed significantly from chance. A linear trend analysis performed on these data was also far from significant [$F(1, 8) = 0.166$].

DISCUSSION

The question addressed by this research is whether K responses reflect high-confident familiarity-based decisions, as is often assumed (and as standard R/K instructions stipulate), or decisions that are based on relatively low-strength memories that involve lesser degrees of recollection. We investigated this issue by using a source memory procedure and found that source recollection accuracy was significantly above chance for items that were declared to be old and known (cf. Hicks, Marsh, & Ritschel, 2002). Moreover, this held true even when the source question was asked before the old/new question (i.e., after participants had often just successfully recollected information about the test item). This is the key contribution of the present study, and it is entirely consistent with a signal-detection interpretation of R/K judgments (Wixted, 2007; Dunn, 2004; Wixted & Stretch, 2004; Donaldson, 1996). The signal-detection account does not necessarily assume that recognition decisions are based on a single process. The dual-process version of that theory assumes that remember judgments reflect memories that are high in both recollection and familiarity, whereas know judgments reflect memories that are lower in both processes (Wixted, 2007).

Dissociations in neural activity between remembering and knowing are often construed as supporting the idea that (a) subjective reports can easily distinguish between recollection and familiarity and (b) different

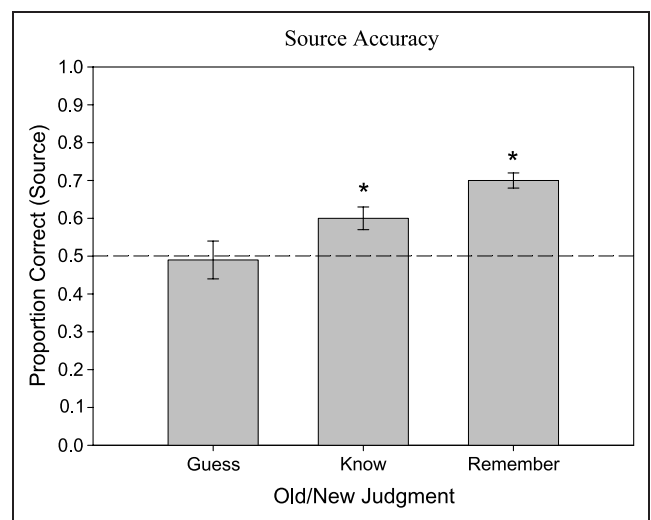


Figure 2. Source accuracy for remember, know, and guess responses collapsed across version 1 and version 2 ($n = 26$). For both remember and know responses, mean proportions with correct source judgments were significantly above the 50% chance level ($p < .01$).

regions of the brain underlie those processes. However, the present results suggest that it may be worth considering what those dissociations would mean if the signal-detection account is correct. In the case of functional magnetic resonance imaging, one simple possibility is that the relationship between neural activity and memory strength (as indexed by misses, K responses, and R responses) is nonlinear and that the nature of that non-linearity differs depending on the brain structure in question. That this might be true should not be surprising given that one class of neurons responds to novelty, whereas another class responds to prior occurrence (Rutishauser, Mamelak, & Schuman, 2006; Viskontas, Knowlton, Steinmetz, & Fried, 2006). The memory strength signal is, presumably, a joint function of the activity of these neurons. If the proportion of novelty-detecting neurons and prior-occurrence-detecting neurons differs across the brain structures of the medial-temporal lobe, then the relationship between memory strength and neural activity is likely to differ as well (perhaps qualitatively). Whatever the reason for the observed dissociations between remembering and knowing, our point is that such dissociations should not be attributed to recollection and familiarity because the R/K procedure probes degrees of recollection, not distinct memory processes.

APPENDIX

Test response instructions for the remember/know paradigm that were read to participants in two experiments which examined process purity (Gardiner & Richardson-Klavehn, 2000):

Recognition memory is associated with two different kinds of awareness. Sometimes when you recognize a word on the test list as one from the first session, recognition will bring back to mind something you remember thinking about when the word appeared then (on the first session list). You recollect something you consciously experienced at that time. In a case like this, click the Remember button. But sometimes recognizing a word as one you saw during the first session will not bring back to mind anything you remember about seeing it then. Instead, the word will seem familiar, so that you feel confident it was the one you saw yesterday, even though you don't recollect anything you experienced when you saw it then. Click the Know button in a case when recognition is accompanied by strong feelings of familiarity in the absence of any recollective experience.

There will also be times when you do not remember the word, nor does it seem familiar, but you might want to guess that it was one of the words you saw during the first session. Click Guess if your response is really just a guess.

Reprint requests should be sent to John T. Wixted, Department of Psychology, 0109, University of California, San Diego, La Jolla, CA 92093-0109, or via e-mail: jwixted@ucsd.edu.

Note

1. These authors ultimately concluded that, in their experiment, R/K judgments were not processed pure and instead reflected different levels of memory strength.

REFERENCES

- Aggleton, J. P., Vann, S. D., Denby, C., Dix, S., Mayes, A. R., Roberts, N., et al. (2005). Sparing of the familiarity component of recognition memory in a patient with hippocampal pathology. *Neuropsychologia*, *43*, 1810–1823.
- Brewer, J. B., Zhao, Z., Desmond, J. E., Glover, G. H., & Gabrieli, J. D. E. (1998). Making memories: Brain activity that predicts how well visual experience will be remembered. *Science*, *281*, 1185–1187.
- Conway, M. A., & Dewhurst, S. A. (1995). Remembering, familiarity, and source monitoring. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology*, *48*, 125–140.
- Curran, T., & Hintzman, D. (1995). Violations of the independence assumption in process dissociation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *21*, 531–547.
- Donaldson, W. (1996). The role of decision processes in remembering and knowing. *Memory & Cognition*, *24*, 523–533.
- Dunn, J. C. (2004). Remember-know: A matter of confidence. *Psychological Review*, *111*, 524–542.
- Eldridge, L. L., Engel, S. A., Zeineh, M. M., Bookheimer, S. Y., & Knowlton, B. J. (2005). A dissociation of encoding and retrieval processes in the human hippocampus. *Journal of Neuroscience*, *25*, 3280–3286.
- Eldridge, L. L., Knowlton, B. J., Furmanski, C. S., Bookheimer, S. Y., & Engel, S. A. (2000). Remembering episodes: A selective role for the hippocampus during retrieval. *Nature Neuroscience*, *3*, 1149–1152.
- Gardiner, J., & Richardson-Klavehn, A. (2000). Remembering and knowing. *The Oxford handbook of memory*. New York: Oxford University Press.
- Gardiner, J. M., Java, R. I., & Richardson-Klavehn, A. (1996). How levels of processing really influence awareness in recognition memory. *Canadian Journal of Experimental Psychology*, *50*, 114–122.
- Gardiner, J. M., Richardson-Klavehn, A., & Ramponi, C. (1997). On reporting recollective experiences and “direct access to memory systems”. *Psychological Science*, *8*, 391–394.
- Gonsalves, B. D., Kahn, I., Curran, T., Norman, K. A., & Wagner, A. D. (2005). Memory strength and repetition suppression: Multimodal imaging of medial temporal cortical contributions to recognition. *Neuron*, *47*, 751–761.
- Gronlund, S. D., Edwards, M. B., & Ohrt, D. D. (1997). Comparison of the retrieval of item versus spatial position information. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *23*, 1261–1274.
- Henson, R. N., Rugg, M. D., Shallice, T., Josephs, O., & Dolan, R. J. (1999). Recollection and familiarity in recognition memory: An event-related fMRI study. *Journal of Neuroscience*, *19*, 3962–3972.
- Hicks, J. L., Marsh, R. L., & Ritschel, L. (2002). The role of recollection and partial information in source monitoring.

- Journal of Experimental Psychology: Learning, Memory, and Cognition*, 28, 503–508.
- Hintzman, D. L., & Caulton, D. A. (1997). Recognition memory and modality judgments: A comparison of retrieval dynamics. *Journal of Memory and Language*, 37, 1–23.
- Holdstock, J. S., Mayes, A. R., Gong, Q. Y., Roberts, N., & Kapur, N. (2005). Item recognition is less impaired than recall and associative recognition in a patient with selective hippocampal damage. *Hippocampus*, 15, 203–215.
- Holdstock, J. S., Mayes, A. R., Roberts, N., Cezayirli, E., Isaac, C. L., O'Reilly, R. C., et al. (2002). Under what conditions is recognition spared relative to recall after selective hippocampal damage in humans? *Hippocampus*, 12, 341–351.
- Johnson, M. K., Kounios, J., & Reeder, J. A. (1994). Time-course studies of reality monitoring and recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 20, 1409–1419.
- Mandler, G. (1980). Recognizing: The judgment of previous occurrence. *Psychological Review*, 87, 252–271.
- Moscovitch, D. A., & McAndrews, M. P. (2002). Material-specific deficits in “remembering” in patients with unilateral temporal lobe epilepsy and excisions. *Neuropsychologia*, 40, 1335–1342.
- Otten, L. (in press). Fragments of a larger whole: Retrieval cues constrain observed neural correlates of memory encoding. *Cerebral Cortex*.
- Perfect, T., Mayes, A. R., Downes, J. J., & van Eijk, R. (1996). Does context discriminate recollection from familiarity in recognition? *Quarterly Journal of Experimental Psychology*, 49A, 797–813.
- Rutishauser, U., Mamelak, A., & Schuman, E. (2006). Single-trial learning of novel stimuli by individual neurons of the human hippocampus–amygdala complex. *Neuron*, 49, 805–813.
- Uncapher, M., & Rugg, M. (2005). Encoding and the durability of episodic memory: A functional magnetic resonance imaging study. *Journal of Neuroscience*, 25, 7260–7267.
- Verfaellie, M., Cook, S. P., & Keane, M. M. (2003). Absence of size congruency effects in amnesic patients' recognition: A failure of perceptually based recollection. *Neuropsychology*, 17, 108–114.
- Viskontas, I., Knowlton, B., Steinmetz, P., & Fried, I. (2006). Differences in mnemonic processing by neurons in the human hippocampus and parahippocampal regions. *Journal of Cognitive Neuroscience*, 84, 1654–1662.
- Wheeler, M., & Buckner, R. (2004). Functional–anatomic correlates of remembering and knowing. *Neuroimage*, 21, 1337–1349.
- 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.
- Woodruff, C. C., Johnson, J. D., Uncapher, M. R., & Rugg, M. D. (2005). Content-specificity of the neural correlates of recollection. *Neuropsychologia*, 43, 1022–1032.
- Yonelinas, A. P., Kroll, N. E., Dobbins, I., Lazzara, M., & Knight, R. T. (1998). Recollection and familiarity deficits in amnesia: Convergence of remember–know, process dissociation, and receiver operating characteristic data. *Neuropsychology*, 12, 323–339.
- Yonelinas, A. P., Kroll, N. E., Quamme, J. R., Lazzara, M. M., Sauve, M. J., Widaman, K. F., et al. (2002). Effects of extensive temporal lobe damage or mild hypoxia on recollection and familiarity. *Nature Neuroscience*, 5, 1236–1241.