Psychology faculty/staff works

By Department

2017

The influence of recollection and familiarity in the formation and updating of associative representations

Jason D. Ozubko
SUNY Geneseo

Morris Moscovitch

Gordon Winocur

Follow this and additional works at: https://knightscholar.geneseo.edu/psychology-faculty

Recommended Citation

Research

The influence of recollection and familiarity in the formation and updating of associative representations

Jason D. Ozubko,[1] Morris Moscovitch,[2] and Gordon Winocur[2]

[1]Department of Psychology, SUNY Geneseo, Geneseo, New York 14454, USA; [2]Rotman Research Institute, Baycrest Centre, Toronto, Ontario M6A 2E1, Canada

Prior representations affect future learning. Little is known, however, about the effects of recollective or familiarity-based representations on such learning. We investigate the ability to reuse or reassociate elements from recollection- and familiarity-based associations to form new associations. Past neuropsychological research suggests that hippocampal, and presumably recollective, representations are more flexible than extra-hippocampal, presumably familiarity-based, representations. We therefore hypothesize that the elements of recollective associations, as opposed to familiarity-based representations, may be more easily manipulated and decoupled from each other, and facilitate the formation of new associations. To investigate this hypothesis we used the AB/AC learning paradigm. Across two recall studies we observed an advantage in learning AC word pairs if AB word pairs were initially recollected. Furthermore, AB word pairs were more likely to intrude during a final AC test if those AB word pairs were initially familiarity-based. A third experiment using a recognition version of the AB/AC paradigm ruled out the possibility that our findings were due to memory strength. Our results support the idea that elements in recollective associative traces may be more discretely coded, leading to their flexible use, whereas elements in familiarity-based associative traces are less flexible.

Dual-process accounts of episodic memory posit that episodic memory can be divided into two qualitatively distinct forms: recollection and familiarity (Mandler 1980; Tulving 1985; Gardiner 1988; Jacoby 1991). Whereas familiarity is characterized as simply having a sense of knowing or acquaintance with stimuli that were experienced, recollection is considered a flexible form of representation that is linked to the context in which the stimuli were presented and enables the individual to relive the past experience associated with the stimuli (see Yonelinas 2002 for a review). Though not a direct relation, researchers have demonstrated that recollection is often supported by the hippocampus whereas familiarity is not (Cohen and Eichenbaum 1993; Eichenbaum et al. 2007; Schacter et al. 2012; see Wixted and Squire 2011 for alternative views about the nature of hippocampal representations).

Neuropsychological evidence also suggests that hippocampal representations are more flexible than representations that rely on extra-hippocampal structures, such as the peri-rhinal and parahippocampal cortices (Cohen and Eichenbaum 1993; Eichenbaum et al. 2007; Schacter et al. 2012). There is also considerable evidence that the flexible use of learned, relational associations depends on the hippocampus (Myers et al. 2003; Eichenbaum et al. 2007; van der Jeugd et al. 2009; see Olsen et al. 2012; see Zeithamova et al. 2012, for reviews). For example, in one fMRI study, Preston et al. (2004) demonstrated that although the posterior hippocampus was activated in response to the retrieval of learned associations, the anterior hippocampus was activated bilaterally when participants were asked to make judgments that required inferences across learned associations (see also Heckers et al. 2004; Zalesak and Heckers 2009). Likewise, Shohamy and Wagner (2009) found that successful generalization from overlapping learned associations to new ones, depends on hippocampal processes involved in integrating the information from learned associations at encoding. Indeed, such findings are pervasive in animal, as well as human, hippocampal research. In one study it was found that although hippocampally damaged rats learned to make odor discriminations between pairs of odors, they were unable to continue to perform the task if the learned odors were rearranged to form novel pairs, supporting the view that the hippocampus is involved with the flexible reorganization of memories based on task demands (Eichenbaum et al. 1989). More recently, researchers have argued that the hippocampus plays a key role in the flexible use of relational information across a number of cognitive and social tasks, including areas as diverse as navigation, exploration, imagination, creativity, character judgments, empathy, social discourse, and language (see Rubin et al. 2014 for review).

Thus, a wide range of research, findings suggest that recollective (or hippocampally dependent representations) are more flexible than familiarity-based representations. Although there have been extensive investigations of the characteristics of recollection-based and familiarity-based representations, and the factors that promote their encoding and retrieval, less work has been devoted to exploring the influence that these two types of memory representations have on later learning (though see Schacter et al. 2012). Understanding how memories, whether they are recollection- or familiarity-based, can support the formation of novel yet overlapping representations may provide important new insights into the nature of memory representation, and the varied roles that these two types of memory play in other domains (e.g., decision-making and problem solving; Myers

et al. 2003; Eichenbaum et al. 2007; van der Jeugd et al. 2009; Olsen et al. 2012; Zeithamova et al. 2012; Moscovitch et al. 2016).

We begin with a discussion of neuropsychological patient work regarding the flexibility of hippocampal-based representations during relearning. Though recollective representations are thought to be hippocampal in nature, even advocates of this view would not argue that recollection is wholly dependent on hippocampal function (Cohen and Eichenbaum 1993; Eichenbaum et al. 2007; Schacter et al. 2012). By examining past neuropsychological findings to develop our hypothesis regarding recollection and relearning, we aim to validate and extend existing neuropsychological accounts of hippocampal function and recollection, while grounding our hypotheses about recollection and familiarity in neuropsychological models of memory representation. The novel hypothesis that we examined is whether the elements of recollective associations, as opposed to familiarity-based representations, are more easily manipulated and decoupled from each other, and facilitate the formation of new associations in neurologically intact individuals. This hypothesis is consistent with cognitive models of recollection, which espouse the idea that recollection-based representations are relational in nature, in that they incorporate both the elements of an experience and their associations to one another (Cohen and Eichenbaum 1993; O'Reilly and Rudy 2001; Olsen et al. 2012). In essence, we ask whether recollective representations are indeed "flexible," as one would predict if they were hippocampally supported, and investigate the consequences of having such representations for learning novel, yet related, and potentially interfering, material.

On the flexibility of hippocampal and recollective representations

Traditionally, the effects of prior memory on subsequent learning have been studied by examining transfer or interference effects from one memory to another. One widely used procedure is AB–AC learning, in which one examines the effects of learning to associate A with B on learning to form a new association of A with C (where A, B, and C are often words). Although there is an extensive literature on this topic (Underwood 1949; Postman 1962), and researchers have used related paradigms to show that recollective representations are resistant to interference (e.g., Jones and Atchley 2006), no studies, to our knowledge, have directly examined the effects of recollection and familiarity-based associations on learning subsequent associations.

Regarding hippocampal representations specifically, several studies examining individuals with hippocampal damage (i.e., amnesic patients), using an AB/AC recall paradigm, have documented that the hippocampus is important for reducing interference between previously learned information and novel information (Winocur and Weiskrantz 1976; Winocur and Kinsbourne 1978; Kinsbourne and Winocur 1980; Winocur et al. 1996). For example, Winocur et al. (1996) demonstrated that patients with left-sided medial temporal lobe (MTL) lesions (including the hippocampus) had difficulty learning AC word pairs after learning AB word pairs, and demonstrated more AB word intrusions during the AC test than did controls or patients with right-sided (i.e., nonverbal hemispheric) lesions. Further, Winocur et al. (1996) demonstrated that similar difficulties in learning AC word pairs were observed in healthy young and old controls when those participants engaged in an implicit memory test, which is not mediated by the hippocampus. Presuming that recollective representations are dependent on the hippocampus (or at least more so than are familiarity-based representations), in these studies patients may have relied more on familiarity than on recollection (Yonelinas 2002), suggesting that familiarity-based memory is indeed more rigid and less flexible than recollection (for similar results see also Winocur and Weiskrantz 1976; Winocur and Kinsbourne 1978; Kinsbourne and Winocur 1980; and see Winocur and Moscovitch 1983 for similar results from older participants).

Consistent with the findings of Winocur et al. (1996), Hay et al. (2002) used process-dissociation procedure (PDP), adapted from Hay and Jacoby (1996), to determine whether the interference resulted from impaired controlled processes (which are often related to recollection-based processes) or automatic processes (which are sometimes related to familiarity-based processes) in patients with MTL lesions. As expected, controlled processes were shown to be impaired, and resulted in greater AB intrusions during the AC test, suggesting that familiarity-based memories may interfere with subsequent learning. If controlled processes are taken as an approximate measure of hippocampal and recollective influences, and automatic processes are taken to reflect extra-hippocampal familiarity-based influences, then these results too are consistent with the idea that recollective representations may be more flexible than familiarity-based representations in terms of forming new associations.

Though patient studies by both Winocur et al. (1996) and Hay et al. (2002) converge on the idea that hippocampal, and thus recollective representations, are more flexible when it comes to later learning, these studies are at best only suggestive. Namely, were recollection or familiarity were not specifically measured in any of these studies, meaning that we must speculate to some degree about what these studies tell us about recollection and familiarity. Research supports the idea that recollection is often hippocampally supported whereas familiarity is not, but as already mentioned, there is some disagreement as to how much recollection is a reflection of hippocampal representation (Wixted and Squire 2011). As well, none of these studies used a baseline condition in which no prior learning occurred. Hence, even if our interpretations regarding recollection and familiarity and hippocampal function are completely accurate, it is unclear whether recollection could have facilitated subsequent learning or merely not interfered with it. Finally, neither of these studies ruled out the possibility that memory strength, rather than recollection versus familiarity, was the critical factor that influenced later learning. Some researchers have pointed out that in many conditions recollections tend to be high-confidence responses whereas familiarity-based responses, on average, tend to be less confident (Donaldson 1996; Wixted and Stretch 2004; Wais et al. 2010). Because of this memory strength difference, it is possible that rather than reflecting differences between recollection and familiarity, past studies are truly reflecting a difference that is being driven by memory strength. Therefore, past patient work supports the idea that recollections may be more flexible than familiarity-based representations, but several important questions remain. In the present study, we will address these issues.

Current study

The present study used the AB–AC paradigm to test directly the hypothesis that recollection-based memories support, and possibly facilitate, the acquisition of related, yet novel, information better than familiarity-based memories. Importantly, this effect could be shown either as an increase in recall and recognition rates for the related material, a decrease in intrusions and false alarms to the originally learned material when testing the related material, or both. In the following experiments, participants first learned a series of AB word pairs (e.g., FIRE–DOG). Subsequently, they learned a series of AC word pairs in which the first word from each pair was shared with a corresponding AB word pair (e.g., FIRE–TREE). By noting which AB word pairs were rated as

recollected versus familiar, we could predict which AC word pairs would be easier to learn. Specifically, based on our hypothesis, the AC word pairs that correspond to recollected AB word pairs should be learned more easily, show less interference than AC word pairs that correspond to nonrecollected (i.e., familiar) AB word pairs, or both. Experiments 1 and 2 address the influence of recollection and familiarity on relearning using AB/AC recall paradigms. In Experiment 3, we adopt a recognition version of the AB/AC paradigm, so as to directly examine the potential role of memory strength in any observed effects.

Experiment 1

Experiment 1 investigated the ability of healthy young adults to learn new AC associations after having acquired AB associations that were either recollected or familiar. Though an extensive literature on the AB–AC paradigm exists (Underwood 1949; Postman 1962; Jones and Atchley 2006), no one has examined whether AC learning is influenced by whether the prior AB associations are recollection-based or familiarity-based. Thus, we combined the cued recall test of the typical AB/AC recall paradigm with remember/know (R/K) style judgments (Tulving 1985) to gauge recollection and familiarity, respectively. Despite concerns as to whether R/K judgments provide proper estimates of recollection and familiarity (e.g., Donaldson 1996; Hirshman and Master 1997; Inoue and Bellezza 1998; Wixted and Stretch 2004; Wixted 2007; Rotello and Zeng 2008), research suggests that they do reflect a qualitative distinction between memories, (e.g., Rajaram 1993; Perfect and Dasgupta 1997; Yonelinas 2002; Eichenbaum et al. 2007; Skinner and Fernandes 2007) that converges with independent measures of recollection and familiarity when instructions on how to make R/K responses are strict (see Yonelinas et al. 1996; Yonelinas 2001; Rotello et al. 2005). Consequently, in all our experiments, strict and detailed instructions were given to ensure that obtained judgments would be reasonable indicators of the presence of recollection and familiarity (see Yonelinas et al. 1996; Yonelinas 2001; Rotello et al. 2005). In Experiment 3, we directly investigate the issue of memory strength as a potential confound.

In terms of combining recall with R/K judgments, few researchers have directly examined the issue of familiarity in cued recall, although there is some evidence that familiarity can contribute even to free recall (Mickes et al. 2013; Sadeh et al. 2015). Several studies have shown that stem and word cues can be effectively completed without conscious recall strategies (Schacter and McGlynn 1989; Hay and Jacoby 1996), even by amnesic patients with severe recollective deficits (Verfaellie et al. 2005). Consistent with this, the few studies that directly investigated recollection and familiarity in cued recall provide evidence of nonrecollective, familiarity-based retrieval (Lindsay and Kelley 1996; Brainerd and Reyna 2010). The broader implication here is that if familiarity is interpreted as a general process that reflects ease or fluency with which information is generated and perceived (e.g., Jacoby et al. 1989), then it seems reasonable to assume that familiarity can support recall in forced-response, cued recall paradigms (as in the present experiments). To reinforce this claim, Experiment 3 will use a recognition paradigm to verify the recall findings regarding familiarity in Experiment 1 and 2.[3]

In Experiment 1, participants learned a list of AB word pairs. They then were given A-items as cues and recalled B-items, for

which they also provided ratings of recollection and familiarity. Performance on this test served as our baseline condition. Next, participants studied a list of AC word pairs. The A-items in this list were the same as those in the AB study list. Learning the association between AC items, therefore, represents learning that requires participants to use the A-word element from the original AB association to form the new AC association. Finally, a second cued recall test was given in which participants once again were presented with A-words as cues but this time were asked to recall only the C-words.

Based on our earlier discussion of neuropsychological findings, we expected that AC word pairs that correspond to recollected AB word pairs would be the easiest to learn, even surpassing initial baseline learning levels. Thus, participants should show higher C-item recall rates, lower B-item intrusion rates, or both for corresponding recollection-based AB pairs than familiarity-based AB pairs.

Results and discussion

In terms of verifying the methods of Experiment 1, a hallmark of recollection-based responding is that it is more accurate than familiarity-based responding, with fewer memory errors (Yonelinas 2002). The mean accuracy of B-item recall for recollected responses was near ceiling ($M = 0.92$, $SE = 0.01$) and significantly greater than the mean accuracy of B-item recall for familiar responses ($M = 0.56$, $SE = 0.05$), $t_{(34)} = 7.10$, $P < 0.01$, $d = 2.43$.

Correct recall rates were calculated separately for the AB Test and AC Test. Correct recall rates in the AB Test represent the probability of a participant recalling (i.e., producing a studied word and identifying it as recollected or familiar) the appropriate B-word to a given A-word cue, and in the AC Test, represent the probability of a participant recalling the appropriate C-word to a given A-word cue.

Critical intrusion rates in the AB Test were calculated as the proportion of C-words produced during the AB Test. Because C-words were never exposed to participants during or before the AB Test, there should be a very low probability of these items being produced. However, the critical intrusion rate in the AB Test should provide a reasonable estimate of the likelihood a participant would produce a critical intrusion purely by chance (i.e., with no influence of memory).

To determine the critical intrusion rate in the AC Test, we calculated the proportion of trials during the AC Test in which participants recalled either the same word as they did in AB Test, or the B-word which was studied in the AB Test. For example, imagine a participant had studied FROG-GREEN during the AB study phase and during the AB Test they recalled SWAMP. During the AC Test, either GREEN or SWAMP would be considered a critical intrusion in response to FROG. Because participants were forced to produce a word on each trial, responses that were identified as guesses were not considered to be intrusions. In all cases then, intrusions were counted only when a participant produced a recollection or familiarity response.

Figure 1A plots the probability of recall and intrusion during the AB and AC Test, respectively. The results of the AB Test are further divided to illustrate the proportion of recollection and familiarity responses given during the AB Test. Significantly more recollection responses were produced than familiarity responses during the AB Test, $t_{(34)} = 4.13$, $P < 0.01$, $d = 1.42$. In terms of recall and intrusion rates, more studied items were recalled in the AB Test than in the AC Test, $t_{(34)} = 3.56$, $P < 0.01$, $d = 1.22$ and more critical intrusions were produced in the AC Test than in the AB Test, $t_{(34)} = 6.65$, $P < 0.01$, $d = 2.28$. Hence, memory was better during the AB Test, likely because the AC Test suffered

---

[3]Although we argue that the concept of familiarity translates readily from recognition to cued recall designs, it is less clear how one would interpret "familiarity"-based responses in free recall. For a discussion of this issue and one interpretation, see Mickes, Seale-Carlisle, and Wixted (2013).
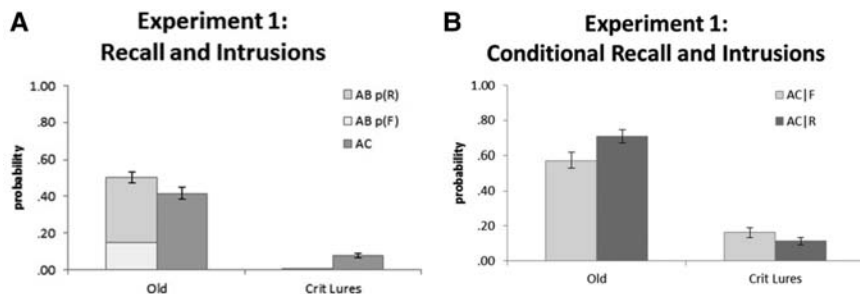
Figure 1. (*A*) Mean proportion of old items and critical lures recalled (i.e., produced and identified as recollected or familiar) during the AB Test and AC Test in Experiment 1. Results of the AB Test are further divided to illustrate the proportion of recollection [p(R)] and familiarity [p(K)] responses. (*B*) Mean proportion of old items and critical lures recalled during the AC Test conditionalized by whether the corresponding AB item was recollected (AC|R) or familiar (AC|F) during the AB Test. In both panels, old items indicate AB pairs in the AB Test and AC pairs in the AC Test, whereas critical lures indicate AC pairs in the AB Test and AB pairs in the AC Test. Error bars represent standard errors of the mean.

proactive interference from AB learning. However, the important question here is whether the proportion of recalls or critical intrusions during the AC Test depended on whether the corresponding AB items were recollected (AC|R) or familiar (AC|F) during the AB Test.

Recalls and critical intrusions during the AC Test, conditionalized on whether the corresponding AB items were recollected or familiar, are shown in Figure 1B. As predicted, there were more recalls for AC|R versus AC|F items, $t_{(34)} = 4.94$, $P < 0.01$, $d = 1.69$, and there were marginally more critical intrusions for the AC|F versus the AC|R items, $t_{(34)} = 2.03$, $P = 0.05$, $d = 0.70$. In fact, using the AB Test as a baseline, the AC|R recall rates were enhanced above baseline, $t_{(34)} = 5.35$, $P < 0.01$, $d = 1.84$, whereas the AC|F recall rates were not, $t_{(34)} = 1.45$, $P = 0.16$, $d = 0.50$. Both the AC|R and AC|F conditions, however, produced more critical intrusions than baseline, with both t's > 5.54, $P < 0.01$, $d = 1.90$.

In sum, it was easier for participants to learn and recall AC items when the corresponding AB items had been recollected, as opposed to when the corresponding AB items had been familiar. Moreover, at least for recall rates, this was not merely a relative difference but a true enhancement above baseline for the AC|R condition.

Experiment 2

In Experiment 1, participants' memory for the AB word pairs was tested before AC learning. One limitation of this approach is that testing itself can act as an encoding event (Roediger and Karpicke 2006). Thus, even if an AB pair is initially only familiar to participants and reported as such during the AB test, the AB items may be reencoded on this test trial as a recollected pair. At least some of the AB items that were classified as familiar may actually have been converted to recollected by the time the AC test occurred. In fact if this were the case, Experiment 1 would have "underestimated" the difference between AC|R and AC|F items, making it a conservative test of our hypothesis. Nonetheless, we conducted Experiment 2, using a modification of the present paradigm which circumvents this issue.

Experiment 2 was similar to Experiment 1 except that instead of AB items being tested before the AC study and test phases, AB items were tested after each AC item was tested, making Experiment 2 a version of modified-modified free recall (MMFR) (Barnes and Underwood 1959). Thus, in Experiment 2, there were only three phases: AB Study, AC Study, and Test. On each test trial, participants were given an A-word cue and asked to recall

the corresponding C-word. Immediately afterward, they were asked to recall the corresponding B-word and then to indicate a rating of recollection, familiarity, or guess for that B-word. In this way, we still obtained recollection/familiarity ratings for AB items in Experiment 2, but not until AC items had been tested, thereby eliminating any potential for AB testing to alter the nature of the AB memory, and resulting in more precise recollection and familiarity ratings.

Results and discussion

In terms of verifying the methods of Experiment 2, the mean accuracy of B-item recall for recollected responses was high ($M = 0.81$, $SE = 0.03$) and significantly greater than the mean accuracy of B-item recall for familiar responses ($M = 0.44$, $SE = 0.04$), $t_{(29)} = 8.95$, $P < 0.01$, $d = 3.32$.

Recalls and intrusions were calculated in the same way as Experiment 1. Namely, recalls occurred when participants produced a B-item during the AB portion of the test or a C-item during the AC portion of the test and identified that item as recollected or familiar. Critical intrusions occurred when participants produced a C-item during the AB portion of the test or a B-item during the AC portion of the test and identified that item as recollected or familiar.

Figure 2A plots the probability of recall and critical intrusion during the AB and AC portion of the test respectively. The AB recall and intrusions are further divided to illustrate the proportion of recollection and familiarity responses given to these items. Significantly more recollection responses were produced than familiar responses for AB items, $t_{(29)} = 3.16$, $P < 0.01$, $d = 1.17$. In terms of recall and critical intrusion rates, more studied items were recalled in the AB Test than in the AC Test, $t_{(29)} = 6.68$, $P < 0.01$, $d = 2.48$, and more critical intrusions were produced in the AC Test than in the AB Test, $t_{(29)} = 5.14$, $P < 0.01$, $d = 1.91$. Memory, therefore, was better for the AB items compared with the AC items, likely due to proactive interference during the AC learning from the AB items. Once again, however, the critical question here is whether the proportion of recalls or critical intrusions during the AC Test depended on whether the corresponding AB items were recollected (AC|R) or familiar (AC|F) during the AB Test.

Recalls and critical intrusions during the AC Test, conditionalized on whether the corresponding AB items were recollected or familiar, are shown in Figure 2B. Replicating Experiment 1, there were more recalls for AC|R versus AC|F items, $t_{(29)} = 4.90$, $P < 0.01$, $d = 1.82$, and there were more critical intrusions for the AC|F versus the AC|R items, $t_{(29)} = 3.29$, $P < 0.01$, $d = 1.22$. Once again, using the AB Test as a baseline, the AC|R recall rates were enhanced above baseline, $t_{(29)} = 5.28$, $P < 0.01$, $d = 1.96$, whereas the AC|F recall rates were not, $t_{(29)} = 0.33$, $P = 0.74$, $d = 0.12$. Both the AC|R and AC|F conditions, however, produced more critical intrusions than baseline, with both t's > 3.23, $P < 0.01$, $d > 1.20$.

The results of Experiment 2 confirm and extend those of Experiment 1. AC learning was facilitated above baseline when AB items were recollected but not when AB items were familiar, whereas the opposite was the case for critical intrusions. Importantly, because the results of Experiment 2 were so similar to those of Experiment 1, we can have confidence in the reliability of these findings despite the different procedures used in these two experiments.
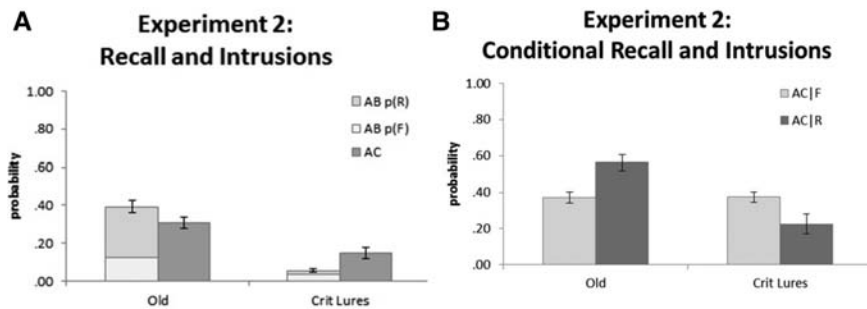
Figure 2. (*A*) Mean proportion of old items and critical lures recalled (i.e., produced and identified as recollected or familiar) during the test in Experiment 2. Results of the AB recall are further divided to illustrate the proportion of recollection [p(R)] and familiarity [p(K)] responses. (*B*) Mean proportion of old items and critical lures recalled during the test conditionalized by whether the corresponding AB item was recollection (AC|R) or familiar (AC|F). In both panels, old items indicate AB pairs for the AB portion of the test and AC pairs for the AC portion of the test, whereas critical lures indicate AC pairs for the AB portion of the test and AB pairs for the AC portion of the test. Error bars represent standard errors of the mean.

Experiment 3

Thus far, we have taken a traditional dual-process approach to the interpretation of recollection and familiarity, presuming that these two measures reflect a qualitative difference in memory representation (see Yonelinas 2002). Single-process accounts offer an alternative view of recollection and familiarity, however, suggesting that the measures do not differ in kind but rather merely in strength (Donaldson 1996; Wixted and Stretch 2004; Wais et al. 2010). From this account, Experiments 1 and 2 could be interpreted as supporting the notion that strong memories reduce interference in AB–AC learning, whereas weaker memories do not. Though we have argued in favor of a dual-process view of recollection and familiarity, the single-process account deserves consideration. Indeed, even if recollection and familiarity are taken to reflect a qualitative distinction in memory, researchers have pointed out that strength is often confounded with recollection and familiarity, such that recollection often represents strong memory, and familiarity represents a range of memory strength, making familiarity, on average, weaker (Wixted and Mickes 2010). Such a model could be termed a "graded dual-process account," and this model too suggests that memory strength serves as an alternative explanation for the data we have observed thus far. Hence, regardless of one's theoretical perspective on dual- versus single-process accounts, it is important to control for strength between recollection and familiarity-based responses when drawing conclusions about differences observed between them.

In Experiment 3, we directly address the issue of memory strength by gathering confidence ratings as a proxy for memory strength. In doing so, we can investigate the influence of strength, and specifically examine whether the strength of representations could explain the effects we observed in Experiments 1 and 2 between recollection and familiarity. If strength plays an important role, then it should determine the pattern of results irrespective of whether the items are recollected or familiar. If strength does not play a primary role, then the recollection and familiarity distinction should determine our results, independent of confidence ratings. Additionally, by equating the confidence between recollected and familiar responses, we can rule out the possibility that our findings are somehow due to item effects, wherein some A items are simply highly memorable and associable. If item effects are driving the findings from Experiments 1 and 2, then high confidence recollected and familiar items should show identical data patterns in Experiment 3.

To examine memory strength, Experiment 3 was designed as an AB–AC recognition task instead of a recall task, and to obtain reliable measures of confidence we included more study and test items than would be feasible in recall. These changes not only offered the opportunity to examine the potential role of memory strength in mediating our effects, but also a chance to extend the findings of Experiment 1 and 2 to recognition. It remains possible that the effects we have reported thus far are somehow artifacts of cued recall itself. This possibility is ruled out if Experiment 3 succeeds in replicating the findings of Experiments 1 and 2. Hence, Experiment 3 serves not only to address the issue of memory strength and item effects, but as a demonstration of the robustness of our findings.

Results and discussion

To verify the recollection/familiarity judgments in the AB Test of Experiment 3, we examined the mean accuracy of these responses. For example, the accuracy of recollection responses was calculated as the probability of a recollection and "old" (confidence rating of 4, 5, or 6) response for an AB item compared with an AB′ or an AD item. Similarly, the accuracy of familiarity responses was calculated as the probability of a familiar and "old" (confidence rating of 4, 5, or 6) response for an AB item compared with an AB′ or an AD item. The mean accuracy of recollection responses ($M = 0.91$, SE $= 0.02$) was significantly greater than the mean accuracy of familiar response ($M = 0.55$, SE $= 0.03$) on the AB Test, $t_{(22)} = 14.33$, $P < 0.01$, $d = 2.56$. Given the success of this manipulation check, we turn now to the critical analyses.

Results were first analyzed by collapsing the six-point confidence ratings into "old" (ratings of 4, 5, or 6) and "new" (ratings of 1, 2, or 3) to calculate hit and false alarm rates (see below for further analyses). Hit rates occurred when participants identified old words as "old" and false alarms occurred when participants identified new words as "old." Hit and false alarm rates for old items, critical lures, and new items for the AB Test and the AC Test are shown in Figure 3A. The AB Test results are further divided by the proportion of items that were identified as recollected and familiar. Hit rates were higher in the AB Test than in the AC Test, $t_{(34)} = 5.69$, $P < 0.01$, $d = 1.95$, indicating better memory for the AB items. False alarms to critical lures in the AB Test and AC Test were not directly comparable here, as in the AB Test they represented false alarms to recombined AB pairs whereas in the AC Test they represented false alarms to previously learned AB items. However, false alarms to critical lures were contrasted against false alarms rates to new items. In both the AB Test and the AC test, false alarm rates to critical lures were greater than false alarm rates to new items, $t_{(34)} = 9.67$, $P < 0.01$, $d = 3.32$ and $t_{(34)} = 6.74$, $P < 0.01$, $d = 2.31$, respectively. Overall however, the more important question was whether the hit rates and false alarms to critical lures in the AC Test varied depending on whether the corresponding AB item was recollected or familiar.

Hit rates and false alarms to critical lures during the AC Test, conditionalized on whether the corresponding AB items were recollected or familiar during the AB Test, are shown in Figure 3B. False alarms to new items are shown for comparison purposes. Interestingly, although false alarms to critical lures were more frequent for AC|F versus AC|R items, $t_{(34)} = 2.64$, $P < 0.05$, $d = 0.91$,
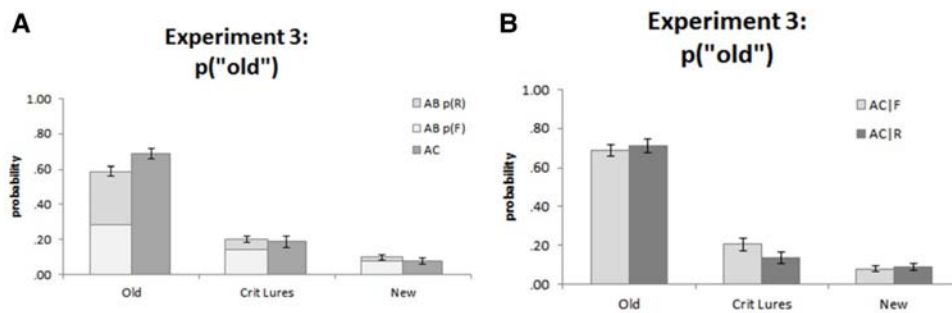
Figure 3. (*A*) Mean proportion of "old" responses to old items, critical lures, and new items in the AB Test and AC Test. Results of the AB Test are further divided to illustrate the proportion of recollection [p(R)] and familiarity [p(F)] responses. (*B*) Mean proportion of "old" responses to old items, critical lures, and new items in the AC Test conditionalized by whether the corresponding AB item was recollected (AC|R) or familiar (AC|F) during the AB Test. In both panels, old items indicate AB pairs in the AB Test and AC pairs in the AC Test, critical lures indicate AB′ pairs in the AB Test and AB pairs in the AC Test, and new items indicate AD pairs in the AB Test and AE pairs in the AC Test. Error bars represent standard errors of the mean.

there was no difference in hit rates between AC|F and AC|R items, $t_{(34)} = 0.99$, $P = 0.33$, $d = 0.34$. Using AB Test hit rates as a baseline, both AC|R and AC|F hit rates were facilitated above baseline, both t's > 4.35, $P < 0.01$, $d > 1.49$. This finding suggests that AC learning is technically facilitated for both AC|R and AC|F items, but the AC representations that are formed are less susceptible to AB interference for AC|R items compared with AC|F items.[4] We will consider the implications of this finding in more detail in the discussion. For now, we continue our analyses by investigating the potential influence of strength. Is it the case that the differences between the AC|R and AC|F conditions could have been due to a difference in strength?

Our preceding analyses were carried out by classifying "old" responses as ratings of 4, 5, or 6 on the six-point confidence scale during the AB and AC Tests. To examine the influence of confidence, the results of the AB Test were recategorized by classifying "old" responses as either high or low confidence. High-confidence "old" responses included ratings of 5 and 6, whereas low confidence "old" responses included ratings of 3 and 4.[5] In our new

analysis, we examined hit and false alarms rates during the AC Test, based on whether the corresponding AB items were recollected or familiar at high or low levels of confidence. It should be noted that no low confidence recollection responses existed and so this category is excluded.

The results of the AC Test conditionalized by both recollection/familiarity and by high and low confidence can be seen in Figure 4A. Importantly, false alarms to critical lures were greater in both the low confidence and high confidence AC|F conditions compared with the (high confidence) AC|R condition, $t_{(34)} = 2.70$, $P < 0.05$, $d = 0.93$ and $t_{(34)} = 3.39$, $P < 0.01$, $d = 1.16$, respectively. There was no difference in false alarm rates to critical lures between the two AC|F conditions, $t_{(34)} = 0.12$, $P = 0.90$, $d = 0.04$. There was similarly no difference between the hit rates of any of the conditions, all t's < 1.59, $P > 0.12$, $d < 0.54$. The observed difference in false alarms to critical lures was not mediated by the strength of representations, as measured by subjective ratings of confidence. Importantly then, the AC|R condition had the lowest false alarm rate to critical lures, independent of the influence of memory strength.

One limitation of examining confidence ratings is that there are typically no low confidence recollection responses, and hence, in our data it was not possible to investigate whether low confidence AC|R items would have a low false alarm rate to critical lures as high confidence AC|R items did. We, therefore, present one more analysis aimed to address the issue of strength. During the AB Test, participants were shown AB and AB′ items. Presumably, if a subject could recognize an AB item as old and reject an AB′ item, the representation that supports that decision must be stronger or more specific than one that would allow them to accept both the AB and AB′ item as old. In essence, a second way to control for the strength of associations is to separate out high accuracy AB items (where AB is accepted and AB′ is rejected) from low accuracy AB items (where both AB and AB′ items are accepted). An added benefit of using this technique is that there were sufficient high accuracy and low accuracy AB items that were recollected for analysis.

Performance during the AC Test, conditionalized on whether AB items were recollected or familiar and whether AB items were high or low accuracy is shown in Figure 4B. False alarms to critical lures were once again greater for AC|F versus AC|R items,

---

[4]We note here that it may be possible that certain AB word pairs were naturally more associable due to their nature (e.g., THE-BOOK would be a highly associable pair), but also by their nature would be easily associable with a new word (e.g., THE-TREE). As a result, is it possible that item-specific effects cause both some AB pairs to be rated as recollected AND to subsequently form strong AC representations (as opposed to a recollected AB pair being the primary cause of the benefit to AC representation)? Though this possibility cannot be ruled out definitively, a supplementary item-analysis was carried out on the data in Experiment 3. In that analysis, we analyzed 415 unique AB word pairs that had been given RKN ratings by at least five different subjects each. We found that only eight of the word pairs were being rated as "recollected" more often than would be expected by chance. In contrast, 16 of the word pairs were rated as "familiar" more often than would be expected by chance, and 20 were rated as "new" more often than would be expected by chance. Dropping these 44 items from our reported analyses had, at best, a minute effect on our results. Reported means and standard errors remained either unchanged or shifted by very small margins (e.g., a mean of 0.64 might have shifted to 0.65 or 0.63, though most means did not). Overall, these supplementary analyses suggest that there were not many AB word pairs that could be considered especially recollectable in general, and furthermore these items did not influence the data patterns that we observed.

[5]Note that although ratings of "3" technically indicated a "new" response, it was the least certain "new" response. That is, a rating of "3" indicates that the subject believes the item might be new, as opposed to a rating of "1" which indicates a subject being certain an item is "new." The logic of treating "3" responses as a low confidence "old" responses comes from receiver operating characteristic (ROC) curve research, which interprets confidence responses as hits and false alarms across varying levels of criteria. For example, in ROC research, it is common to create hits and false alarms that include confidence responses from "2" to "6" together indicating a subject's more lenient (theoretical) criterion for a hit or false alarm. It should also be noted that none of

the reported patterns depend on the inclusion of "3" ratings, and that including both "3" and "4" together to indicate weak hits only serves to produce more reliable estimates across subjects.
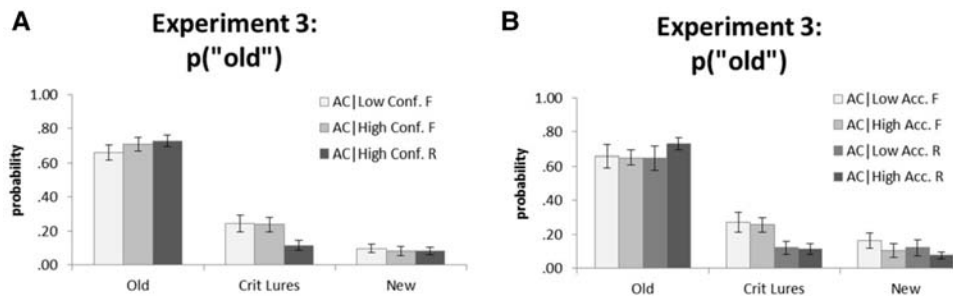
Figure 4. (A) Mean proportion of "old" responses to old items, critical lures, and new items in the AC Test conditionalized by whether the corresponding AB item was high confidence recollected, high confidence familiar, or low confidence familiar during the AB Test. (B) Mean proportion of "old" responses to old items, critical lures, and new items in the AC Test conditionalized by whether the corresponding AB item was high accuracy recollected, low accuracy recollected, high accuracy familiar, or low accuracy familiar during the AB Test. In both panels, old items indicate AC pairs in the AC Test, critical lures indicate AB pairs in the AC Test, and new items indicate AE pairs in the AC Test. Error bars represent standard errors of the mean.

regardless of strength. Namely, for low accuracy items, false alarms to critical lures were greater for AC|F versus AC|R items, $t_{(34)} = 3.44$, $P < 0.01$, $d = 1.18$, and that was true as well for high accuracy items, $t_{(34)} = 3.78$, $P < 0.01$, $d = 1.30$. There was no difference between false alarms to critical lures for high and low accuracy AC|F items, $t_{(34)} = 0.19$, $P = 0.85$, $d = 0.07$, nor for high and low accuracy AC|R items, $t_{(34)} = 1.13$, $P = 0.29$, $d = 0.39$. Interestingly, high versus low accuracy did influence hit rates, but this difference was only significant between high-confidence and low-confidence AC|R items, $t_{(34)} = 10.12$, $P < 0.01$, $d = 3.47$. All other comparisons between hit rates were marginal at best, all t's < 1.81, $P > 0.07$, $d < 0.62$. Although this finding provides some modest evidence that stronger AB representations may facilitate the acquisition of AC items, this finding was found within the AC|R items. That is, there was no evidence that strength influenced AC|F items. Given that strength often varies for familiar items whereas recollected items are typically always quite strong, this finding should quell any concerns that the preceding differences between AC|R and AC|F conditions could have been explained by strength. If anything, strength may play a small role in enhancing the effects observed for AC|R items, but not AC|F items.

The goal of Experiment 3 was to confirm and extend the results of Experiments 1 and 2, using a recognition paradigm. In Experiments 1 and 2, it was found that if an AB item were recollected, it was subsequently easier to learn the corresponding AC item. At the final test, this effect manifested as an increased recall of AC and a decreased likelihood of an AB intrusion. The results of Experiment 3 generally support Experiments 1 and 2, while suggesting that AC learning may be facilitated for both AC|R and AC|F items, but the AC representations that are formed are less susceptible to AB interference for AC|R items compared with AC|F items. This issue is discussed more in the General Discussion.

As a final point, a critic might argue that in Experiment 3 participants may have relied heavily on strategic rejection, which could account for the data pattern observed. For example, during the AC Test if a participant is sure that they studied AC, they can identify that AC item as studied and reject all other items as "new," without evaluating them much. Although we cannot rule out this possibility definitively, if participants were using such a strategy, we might expect that the confidence with which they can identify items would influence their false alarm rates. Namely, when participants are very sure of their memory for AC, they could accept AC and reject all others, but when they are less sure they cannot categorically reject all non-AC items, and hence, false alarms should increase. We saw no such rise in intrusions linked to confidence in our data, suggesting that this type of strategy was not playing a large role in our results. As

well, Experiment 3 replicated the findings of Experiments 1 and 2, which used a completely different procedure, lending further support to the idea that the results of Experiment 3 were dependent primarily on recollection and familiarity, rather than other idiosyncrasies related to its design.

General discussion

Neuropsychological evidence suggests that hippocampal memories are more flexible than extra-hippocampal memories, but the implications of this idea are only beginning to be explored. In the present study, we investigated the novel hypothesis that recollection-based and familiarity-based associative memories, by being differentially supported by the hippocampus, should contribute differently to the formation of new associations in neurologically intact individuals. Although we reviewed some patient findings which are consistent with this hypothesis (Winocur and Weiskrantz 1976; Winocur and Kinsbourne 1978; Kinsbourne and Winocur 1980; Winocur et al. 1996; Hay et al. 2002), it has never been tested directly. Across three experiments, using the AB/AC learning paradigm, we found converging evidence that AC learning was facilitated above baseline when AB pairs were recollected. In recall conditions, AC learning was less affected when AB pairs were familiar, although using a more sensitive recognition paradigm we showed that AC learning may indeed be facilitated when AB pairs are familiar. However, the manner in which recollected and familiarity-based representations support AC learning is fundamentally different: B-item false memories (intrusions and false alarms) during the AC test were greater when AB pairs were familiar when compared with when they were recollected. Thus, recollection-based memories not only support the formation of new associations but do so in a manner that, compared with familiarity-based memories, minimizes interference, regardless of the strength of the recollection-based or familiarity-based memory.

Regarding Experiment 3 specifically, when AB representations are recollected, they supported the acquisition of AC items. This was shown by an increased hit rate to AC items above baseline, and a decreased false alarm rate to AB items compared with when AB representations were familiar. AC learning, however, was also supported when AB representations were familiar, with hit rates to AC items being above baseline, suggesting learning was facilitated. However, false alarms to AB items during the AC Test were higher when AB representations were familiarity-based, than when they were recollected. Together, these results suggest that recollected AB representations support AC learning by facilitating the acquisition of AC items in a way that simultaneously reduces interference with existing AB representations.

Familiarity-based AB representations also seem to support AC learning by facilitating the acquisition of AC items to a degree, but the AC representations that are formed are not independent of the previously acquired AB representations. In essence, a familiarity-based AB representation can facilitate AC learning, but it does so in such a manner that the AB and AC representations are overlapping and interfere with one another during the AC Test. In recall tasks such as in Experiment 1 and 2, because both AC and AB are competing for retrieval during the AC Test, we would have expected recall rates for AC items to be suppressed by the occasional intrusion of AB items. That is why there was no facilitative effect in recall rates for AC|F items in Experiments 1 and 2.

Overall, Experiment 3 shows effects that were similar and consistent to those observed in Experiments 1 and 2 on cued recall, affirming our supposition that familiarity, as measured in a cued recall paradigm, would be analogous to that observed in recognition. As well, using two techniques to address the issue of strength, we found little evidence that strength was the driving force behind our observed effects. Instead, the differences that arose were dependent on recollection and familiarity. Although both recollective and familiarity-based representations may support AC learning to a degree, recollective representations do so in a manner that minimizes interference in comparison to familiarity-based representations.

Neuropsychological implications

The present findings connect with emerging work which demonstrates that hippocampally dependent representations, which presumably are recollective, support the flexible use of existing relational associations to draw inferences (Myers et al. 2003; Eichenbaum et al. 2007; van der Jeugd et al. 2009; Olsen et al. 2012; Zeithamova et al. 2012; Moscovitch et al. 2016). Specifically, the present work shows that recollective representations are flexible in that the individual elements that comprise the representation can be selectively reassociated during later learning. This principle may underlie the operations implicated in performing some tests of inferential reasoning. For example, in an fMRI study, Preston et al. (2004) demonstrated that although the posterior hippocampus was activated in response to the retrieval of learned associations, the anterior hippocampus was activated bilaterally when participants were asked to make judgments that required inferences across learned associations (see also Heckers et al. 2004; Zalesak and Heckers 2009). Likewise, Shohamy and Wagner (2009) found that successful generalization from overlapping learned associations to new ones, depends on hippocampal processes involved in integrating the information from learned associations at encoding. In a subsequent study, Foerde et al. (2013) showed that similar, hippocampally dependent processes were operating in transferring the value of one item, to a second item with which it was experimentally associated. Underlying these, and similar phenomena (for reviews, see Olsen et al. 2012; Zeithamova et al. 2012), is the finding that the hippocampus supports flexible combination or recombination of elements in existing associations to form new associations.

If we accept that recollective/hippocampal representations are flexible, one could ask why familiarity-based representations are not. One suggestion is that they are unitized, a process whereby paired elements, such as the words FIRE and DOG, are represented as a single indivisible unit, FIREDOG (Laberge and Samuels 1974; Graf and Schacter 1989; Giovanello et al. 2006; Quamme et al. 2007). Quamme et al. (2007) noted that, whereas relational associations are believed to rely on hippocampal representation, unitized associations may be akin to individual items and dependent on areas that represent familiarity, such as the perirhinal cortex (Aggleton and Brown 1999; Yonelinas 2002; Bowles et al. 2007; Mayes et al. 2007). To test this hypothesis, Quamme et al. had patients with hippocampal damage and clear recollective deficits learn pairs of words under either associative encoding conditions (i.e., two words were presented in a sentence) or unitized encoding conditions (i.e., two words were presented as a compound word with a definition for that compound word below). As expected, the patients showed typical associative memory impairments when word pairs were learned under associative encoding conditions, but not when word pairs were unitized at encoding. In related work, researchers have used ERPs (Rhodes and Donaldson 2007; Diana et al. 2011) and receiver operating characteristic (ROC) curves (Diana et al. 2008) to demonstrate that familiarity can support associative source judgments when source and item-information is unitized. Thus, the unitization of associations allowed individuals to represent associations with familiarity, instead of recollection (see also Yonelinas et al. 1999; Giovanello et al. 2006; Bastin et al. 2010; Tibon et al. 2012).

Although unitization is beneficial in some circumstances, inasmuch as it may help individuals form associative representations using familiarity-based processes, a drawback to unitization is that it may be difficult to re-form or change associations. That is, in contrast to recollection-based memories which may code individual elements and the associations between elements, unitized representations seem to integrate elements and their associations into a gestalt or indivisible entity. Consistent with this account, research has shown that unitizing items can lead to a memory benefit for the combined pair (i.e., the association), while at the same time impair memory for the individual items themselves (i.e., item-memory; Pilgrim et al. 2012; Murray and Kensinger 2013). If those elements or associations need to be changed at a later time, it may be difficult either to effect such a change in unitized representations, or to access the representations of the individual items to be used as a component in a new association. Therefore, individuals may be more susceptible to interference from the previously encoded unitized association, because of the difficulty in separating the elements from the previous association to form an independent, novel association with some of the elements. As a result, given an element of the new association as a cue, participants would be more likely to intrude the previous (unitized) association, as we observed in our experiments. In sum, the current results are consistent with the notion that familiarity-based associations may sometimes be unitized, and when this occurs, the unitized associations offer less benefit to later learning of novel, overlapping information, because they may act to compete at a final test, and increase false memories.

Conclusion

In three experiments we have shown that recollective associative memories demonstrate more flexibility than familiarity-based associative memories. Participants appeared to use elements from recollective associative memories to facilitate the formation of new associative memories more easily than elements from familiarity-based associative memories. Assuming that recollection is supported by the hippocampus whereas familiarity, typically, is not, these findings are consistent with neuropsychological evidence that shows hippocampal representations are more flexible than extra-hippocampal representations (Cohen and Eichenbaum 1993; Eichenbaum et al. 2007; Schacter et al. 2012). These results also fit with cognitive accounts that suggest the elements of recollection are discretely coded and bound together in such a manner that the elements can easily be individually accessed for later learning (Cohen and Eichenbaum 1993; O'Reilly and Rudy 2001), whereas familiarity-based traces are less flexible (Olsen et al. 2012; Zeithamova et al.

2012), possibly because they are unitized , making it difficult to extract individual elements for reuse (Giovanello et al. 2006; Quamme et al. 2007).

Materials and Methods

Experiment 1

Method

Participants. Thirty-nine undergraduate students from the University of Toronto Scarborough, completed Experiment 1 online in exchange for course credit. Four participants produced either no recollection or familiarity responses during the AB Test. Because AC Test results would be binned by recollection and familiarity responses during the AB Test, these participants were dropped. Thus, a total of 35 participants were included in the analyses. All recruitment and testing followed the appropriate ethical guidelines for the University of Toronto.

Materials. A word pool of 155 cue-target sets was created from the free association norms of Nelson et al. (2004). This pool consisted of sets of three words: one cue and two related targets. The mean probability that a cue would give rise to a target was 0.52 (SD = 0.15). For each participant, 40 cue-target sets were randomly selected to be used. In all cases, the cue was designed to be the A-word. In half the cases the stronger associate of the cue was designated to be the B-word and the weaker associate of the cue was designated to be the C-word. In the other cases, B-words were the weaker associate of the cue and C-words were the stronger associate.

Finally, a pool of 543 additional cues was selected from the free association norms of Nelson et al. (2004). These cues were unrelated to each other and to any of the 155 cue-target sets that were previously selected. Thus, these 543 words served as a random list of unrelated words. For each participant, 40 cue-target sets were randomly created from this pool of 543 words. Each cue-target set was created by randomly selecting one word to serve as the cue (or A-word), and two other words to serve as the targets (or B- and C-words).

Past work has demonstrated that related word pairs are better represented by familiarity than are unrelated word pairs (Giovanello et al. 2006; Greve et al. 2007, 2011). Additionally, words that are deeply processed rely on recollection more so than words that are shallowly processed (see Craik and Lockhart 1972). By having related and unrelated word pairs, some of which will be processed deeply and others shallowly (see Procedure below), our goal was to ensure a distribution of recollection and familiarity-based memory for the pairs in our experiment. Initial analyses of the AB test data confirm that these manipulations had effects consistent with the past literature: Overall familiarity, as measured using the independent remember-know procedure (see Yonelinas and Jacoby 1995; Jacoby et al. 1997; Ochsner 2000; Mangels et al. 2001)),[6] was greater for related word pairs (M = 0.29, SE = 0.05) than for unrelated word pairs (M = 0.15, SE = 0.04), $t_{(29)} = 5.02$, $P < 0.01$, $d = 0.65$, and deeply encoded words showed a greater proportion of recollection responses (M = 0.47, SE = 0.04) than did shallowly encoded words (M = 0.21, SE = 0.03), $t_{(29)} = 8.74$, $P < 0.01$, $d = 1.47$.

Procedure. Experiment 1 consisted of four phases: Study 1 (AB Study), Test 1 (AB Test), Study 2 (AC Study), and Test 2 (AC Test). These four phases were run four times per participant. That is, after completing the first AC Test, participants began again with a new AB Study through to a new AC Test. This was repeated until four full runs of the phases had been completed.

During each run, 10 of the 40 related cue-target sets and 10 of the 40 unrelated cue-target sets were selected to be used as stimuli. Thus, in Study 1, participants studied 20 AB word pairs, 10 of which were related and 10 of which were unrelated. The order of word pairs was randomized at the beginning of each phase. Cue-target sets never repeated between runs so that all 40 related and 40 unrelated cue-target sets for each participant were used by the end of the experiment.

During AB Study, AB word pairs were presented individually on the screen. For each word pair, participants either had to indicate which of the two words had more vowels, or whether the words were related. When judging vowels, participants could indicate either "the left word," "equal," or "the right word." When judging whether two words were related participants could indicate that the words were "not related," "somewhat related," or "very related." These tasks acted as shallow and deep encoding manipulations, respectively (see Craik and Lockhart 1972), to ensure there would be enough encoding variability to yield both recollection-based and familiarity-based word pairs. In both cases, participants made their selections by clicking on an option with the mouse and then clicking a NEXT button to proceed to the next trial. There was a 500-ms inter-stimulus interval between trials.

During the AB Test participants were shown A-words individually on the screen and had to recall the corresponding B-word. Participants were told to guess if no word came to mind. After typing in their response, participants were asked why they chose that particular response, and could indicate that they recollected the word, the word was familiar, or they were simply guessing. To ensure the accuracy of recollection and familiarity responses and to ensure they were not confounded with confidence, a detailed set of instructions explaining the difference between recollection and familiarity and confidence were developed for our experiment. In our experiment, "reexperience" and "familiar" were selected as labels for recollection and familiarity in lieu of "remember" and "know." Piloting showed that "reexperience" and "familiar" were more intuitive for participants, and better fit with the descriptions of recollection and familiarity.

Because our experiment was conducted online, in-laboratory piloting was used to confirm that participants understood the written recollection/familiarity instructions independently, without interaction with a researcher. Strict instructions were provided, to ensure that participants did not confuse these concepts with confidence. Participants were also warned that at the end of the experiment they may be prompted with some items that they said they recollected and asked to justify their responses by explaining what specifically they had recollected.

During AC Study, AC word pairs were presented individually on the screen. 4500 msec after a pair was presented, a NEXT button appeared and participants could click it with the mouse to proceed. Participants were instructed to study the AC word pairs for a later memory test. No specific encoding instructions were given for the AC Study phase. There was a 500-msec inter-stimulus interval between trials.

Finally, during AC Test, memory for the AC associations was tested. On each trial, an A-word was shown. Participants were first asked to recall the C-word that corresponded to this A-word, from the study phase. Participants typed in their response and pressed ENTER to proceed. Participants were instructed to guess if they could not recall the C-word, but not to produce B-words.

Experiment 2

Method

Participants. Thirty-eight participants from the same pool as Experiment 1 participated in Experiment 2. Eight participants produced either no recollection or no familiarity responses and were dropped. Thus, a total of 30 participants were included in the analyses.

Materials. The same materials were used as in Experiment 1.

Procedure. The procedure of Experiment 2 was identical to that of Experiment 1 except that instead of a two test phases there

---

[6]When comparing overall levels of familiarity between conditions, examining raw "know" responses would be misleading as the proportion of items recognized and "remembered" differed between these two conditions. Because of these circumstances, the independent remember-know procedure was adopted to allow for a more unbiased comparison of the overall levels of familiarity of the two stimuli sets.
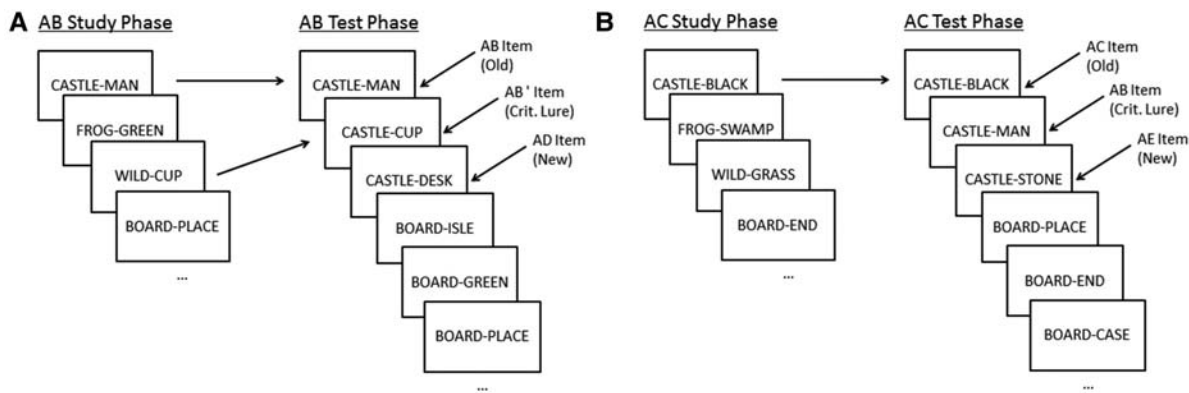
Figure 5. Illustration of the AB Study and Test phases (*A*) and the AC Study and Test phases (*B*) in Experiment 3. During the AB Study phase, participants study pairs of words identified as AB pairs. Afterward, some AB pairs are recombined to form AB′ items. During the AB Test, participants are shown triads of AB, AB′, and AD (new) items and must rate confidence and make recollection/familiarity judgments for each item. During the AC Study phase A items are recombined with novel words to form AC pairs. During the AC Test phase participants are shown triads of AC, AB, and AE (new) items. Hit rates in the AC Test therefore reflect the ability to learn AC items whereas false alarms to critical lures in the AC Test reflect the tendency of previously learned items (i.e., AB items) to interfere during AC retrieval.

was only one. Thus, the three phases of Experiment 2 were Study 1 (AB Study), Study 2 (AC Study), and Test. The Test phase was a version of MMFR. On each test trial, participants were given an A-word cue and asked to recall the corresponding C-word first. Afterward, participants were asked to recall the corresponding B-word and then to indicate how they recalled the B-word by indicating recollection, familiarity, or guess. Thus, although there was a single test phase, AB and AC memory were still probed during this test. For consistency with Experiment 1, we will still refer to the AB Test and AC Test results.

Experiment 3

Method

Participants. Thirty-eight participants from the same pool as Experiments 1 and 2 participated in Experiment 3. Three participants produced either no recollection or familiarity responses during the AB Test and were dropped. Thus, a total of 35 participants were included in the analyses.

Materials. A word pool of 1236 nouns was selected from the MRC Linguistic Database (Wilson 1988). Words were between 5 and 10 letters long ($M = 6.82$, $SD = 1.57$) and had a Kucera–Francis word frequency (Kucera and Francis 1967) between 20 and 395 ($M = 70.76$, $SD = 64.03$). For each participant, 600 words were randomly selected from this pool. These words were randomly divided into sets of 120 A-words, B-words, C-words, D-words, and E-words. From these subsets 120 AB, 60 AC, 60 AD, and 60 AE pairs were arranged.

Procedure. The procedure of Experiment 3 was similar to Experiment 1 except that participants were tested with yes/no recognition instead of cued recall. Furthermore, the experiment was not broken up into runs because performance on recognition was expected to be good even for a large study list. Thus, Experiment 3 consisted of four phases: Study 1 (AB Study), Test 1 (AB Test), Study 2 (AC Study), and Test 2 (AC Test). For illustrative purposes, the different phases and types of stimuli in Experiment 3 are presented in 3.

During the AB Study phase participants studied 120 AB word pairs, which were presented individually for 4000 msec with a 1000-msec inter-stimulus interval. Unlike the previous experiments, participants were given no specific encoding instructions. During the AB Test phase, participants saw 60 studied AB pairs mixed in with 120 new pairs. Half the new pairs were AB′ items, such that the B-item in the pair was taken from another pair that had been studied but was not tested (i.e., the A- and B-words appeared in the study list but not

together).[7] The remaining new items were AD items, such that the A-word had been studied but the D-word was completely novel.

The order of the test list was randomized except that AB, AB′, and AD pairs were grouped together by A-word. Thus, for example, if the AB, AB′, and AD pairs were CASTLE-MAN, CASTLE-CUP, and CASTLE-DESK, respectively (see Fig. 5), these three word pairs would appear sequentially in the test list. The order of which pair would appear first, second, and third was randomized, as was the location of the triad in the overall test list, but all words sharing the same A-word appeared together. This grouping by A-word was done to prevent subsequent memory effects from arising as might occur if a participant was tested on the AD pair or AB′ pair early in the test list which then interfered with original memory for the AB pair which would be tested later in the list. To ensure that ratings of AB, AB′, and AD pairs were as independent as possible, participants were given detailed instructions about the need to evaluate each pair separately, and they were specifically told that they could conceivably rate all pairs as studied or all pairs as new if their feelings or memories supported such claims. That is, participants were told not to rely on strategies of disregarding other items of an A-set once they had spotted an item they thought was studied, and instead to continue to evaluate each item on its own merit.

On each test trial a word pair was presented to participants and they had to judge whether the pair had been studied or was new. Participants made their response on a six-point confidence scale from 6 (sure studied) to 1 (sure new). After making their decision, participants indicated the subjective state of memory on which their recognition decision was based. Here, participants could indicate recollection, familiarity, or guess.

After the AB Test, participants engaged in the AC Study phase. Here, AC word pairs were shown individually for 4000 msec with a 1000-msec inter-stimulus interval. Participants studied 60 AC word pairs which corresponded to the 60 intact AB pairs that were tested in the AB Test (see Fig. 5). Once again, participants were given no specific encoding instruction for these items. During the AC Test, participants saw the 60 studied AC pairs inter-mixed with 120 new pairs. Half of the new pairs were AB items from the AB study and test phase. The remaining new

---

[7]The inclusion of AB′ items was to ensure that participants would need to rely on the associative memory of the AB items when recognizing items and making their R/K judgments. Without AB′ items, participants could presumably just rely on item-recognition (i.e., B versus D items) and ignore the associative information.

pairs were AE pairs (i.e., familiar A-words paired with completely new E-words). Participants provided confidence ratings and recollection/familiarity responses during the AC Test, in the same manner as in the AB Test. As well, although the order of the test list was randomized, AC, AB, and AE pairs were grouped by A-words as in the AB Test.[8]

Acknowledgments

References

Aggleton JP, Brown MW. 1999. Episodic memory, amnesia, and the hippocampal–anterior thalamic axis. *Behav Brain Sci* **22:** 425–444; discussion 444-89.

Barnes JM, Underwood BJ. 1959. "Fate" of first-list associations in transfer theory. *J Exp Psychol* **58:** 97–105.

Bastin C, van der Linden M, Schnakers C, Montaldi D, Mayes AR. 2010. The contribution of familiarity to within- and between-domain associative recognition memory: use of a modified remember/know procedure. *Eur J Cogn Psychol* **22:** 922–943.

Bowles B, Crupi C, Mirsattari SM, Pigott SE, Parrent AG, Pruessner JC, Yonelinas AP, Köhler S. 2007. Impaired familiarity with preserved recollection after anterior temporal-lobe resection that spares the hippocampus. *Proc Natl Acad Sci* **104:** 16382–16387.

Brainerd CJ, Reyna VF. 2010. Recollective and nonrecollective recall. *J Mem Lang* **63:** 425–445.

Cohen NJ, Eichenbaum H. 1993. *Memory, amnesia, and the hippocampal system.* MIT Press, Cambridge, MA.

Craik FIM, Lockhart RS. 1972. Levels of processing: a framework for memory research. *J Verbal Learning Verbal Behav* **11:** 671–684.

Diana RA, Yonelinas AP, Ranganath C. 2008. The effects of unitization on familiarity-based source memory: testing a behavioral prediction derived from neuroimaging data. *J Exp Psychol Learn Mem Cogn* **34:** 730–740.

Diana RA, Van den Boom W, Yonelinas AP, Ranganath C. 2011. ERP correlates of source memory: unitized source information increases familiarity-based retrieval. *Brain Res* **1367:** 278–286.

Donaldson W. 1996. The role of decision processes in remembering and knowing. *Mem Cognit* **24:** 523–533.

Eichenbaum H, Mathews P, Cohen NJ. 1989. Further studies of hippocampal representation during odor discrimination learning. *Behav Neurosci* **103:** 1207–1216.

Eichenbaum H, Yonelinas AP, Ranganath C. 2007. The medial temporal lobe and recognition memory. *Annu Rev Neurosci* **30:** 123–152.

Foerde K, Race E, Verfaellie M, Shohamy D. 2013. A role for the medial temporal lobe in feedback-driven learning: evidence from amnesia. *J Neurosci* **33:** 5698–5704.

Gardiner JM. 1988. Functional aspects of recollective experience. *Mem Cognit* **16:** 309–313.

Giovanello KS, Keane MM, Verfaellie M. 2006. The contribution of familiarity to associative memory in amnesia. *Neuropsychologia* **44:** 1859–1865.

Graf P, Schacter DL. 1989. Unitization and grouping mediate dissociations in memory for new associations. *J Exp Psychol Learn Mem Cogn* **15:** 930–940.

Greve A, van Rossum MCW, Donaldson DI. 2007. Investigating the functional interaction between semantic and episodic memory: convergent behavioral and electrophysiological evidence for the role of familiarity. *Neuroimage* **34:** 801–814.

Greve A, Evans CJ, Graham KS, Wilding EL. 2011. Functional specialisation in the hippocampus and perirhinal cortex during the encoding of verbal associations. *Neuropsychologia* **49:** 2746–2754.

Hay JF, Jacoby LL. 1996. Separating habit and recollection: memory slips, process dissociations, and probability matching. *J Exp Psychol Learn Mem Cogn* **22:** 1323–1335.

Hay JF, Moscovitch M, Levine B. 2002. Dissociating habit and recollection: evidence from Parkinson's disease, amnesia and focal lesion patients. *Neuropsychologia* **40:** 1324–1334.

Heckers S, Zalesak M, Weiss AP, Ditman T, Titone D. 2004. Hippocampal activation during transitive inference in humans. *Hippocampus* **14:** 153–162.

Hirshman E, Master S. 1997. Modeling the conscious correlates of recognition memory: reflections on the remember-know paradigm. *Mem Cognit* **25:** 345–351.

Inoue C, Bellezza FS. 1998. The detection model of recognition using know and remember judgments. *Mem Cognit* **26:** 299–308.

Jacoby LL. 1991. A process dissociation framework: separating automatic from intentional uses of memory. *J Mem Lang* **30:** 513–541.

Jacoby LL, Kelley CM, Dywan J. 1989. Memory attributions. In *Varieties of memory and consciousness: essays in honour of Endel Tulving* (ed. Roediger HLI, Craik FIM), pp. 391–422. Lawrence Erlbaum Associates, Inc., Hillsdale, NJ, England.

Jacoby LL, Yonelinas AP, Jennings JM. 1997. The relation between conscious and unconscious (automatic) influences: a declaration of independence. In *Scientific approaches to consciousness*, pp. 13–47. Erlbaum, Mahwah, NJ.

Jones TC, Atchley P. 2006. Conjunction errors, recollection-based rejections, and forgetting in a continuous recognition task. *J Exp Learn Mem Cogn* **32:** 70–78.

Kinsbourne M, Winocur G. 1980. Response competition and interference effects in paired-associate learning by Korsakoff Amnesics. *Neuropsychologia* **18:** 541–548.

Kucera F, Francis W. 1967. *Computational analysis of present-day American English.* Brown University Press, Providence, RI.

Laberge D, Samuels SJ. 1974. Toward a theory of automatic information processing in reading. *Cogn Psychol* **6:** 293–323.

Lindsay SD, Kelley CM. 1996. Creating illusions of familiarity in a cued recall remember/know paradigm. *J Mem Lang* **35:** 197–211.

Mandler G. 1980. Recognizing: the judgment of previous occurrence. *Psychol Rev* **87:** 252–271.

Mangels JA, Picton TW, Craik FI. 2001. Attention and successful episodic encoding: an event-related potential study. *Brain Res Cogn Brain Res* **11:** 77–95.

Mayes A, Montaldi D, Migo E. 2007. Associative memory and the medial temporal lobes. *Trends Cogn Sci* **11:** 126–135.

Mickes L, Seale-Carlisle TM, Wixted JT. 2013. Rethinking familiarity: remember/know judgments in free recall. *J Mem Lang* **68:** 333–349.

Moscovitch M, Cabeza R, Winocur G, Nadel L. 2016. Episodic memory and beyond: the hippocampus and neocortex in transformation. *Annu Rev Psychol* **67:** 105–134.

Murray BD, Kensinger EA. 2013. A review of the neural and behavioral consequences for unitizing emotional and neutral information. *Front Behav Neurosci* **7:** 42.

Myers CE, Shohamy D, Gluck MA, Grossman S, Kluger A, Ferris S, Golomb J, Schnirman G, Schwartz R. 2003. Dissociating hippocampal versus basal ganglia contributions to learning and transfer. *J Cogn Neurosci* **15:** 185–193.

Nelson DL, McEvoy CL, Schreiber TA. 2004. The University of South Florida free association, rhyme, and word fragment norms. *Behav Res Methods Instrum Comput* **36:** 402–407.

Ochsner KN. 2000. Are affective events richly recollected or simply familiar? The experience and process of recognizing feelings past. *J Exp Psychol Gen* **129:** 242–261.

Olsen RK, Moses SN, Riggs L, Ryan JD. 2012. The hippocampus supports multiple cognitive processes through relational binding and comparison. *Front Hum Neurosci* **6:** 146.

O'Reilly RC, Rudy JW. 2001. Conjunctive representations in learning and memory: principles of cortical and hippocampal function. *Psychol Rev* **108:** 311–345.

Perfect TJ, Dasgupta ZR. 1997. What underlies the deficit in reported recollective experience in old age? *Mem Cognit* **25:** 849–858.

Pilgrim LK, Murray JG, Donaldson DI. 2012. Characterizing episodic memory retrieval: electrophysiological evidence for diminished familiarity following unitization. *J Cogn Neurosci* **24:** 1671–1681.

Postman L. 1962. Transfer of training as a function of experimental paradigm and degree of first-list learning. *J Verbal Learn Verbal Behav* **1:** 109–118.

Preston AR, Shrager Y, Dudukovic NM, Gabrieli JDE. 2004. Hippocampal contribution to the novel use of relational information in declarative memory. *Hippocampus* **14:** 148–152.

Quamme JR, Yonelinas AP, Norman KA. 2007. Effect of unitization on associative recognition in amnesia. *Hippocampus* **17:** 192–200.

Rajaram S. 1993. Remembering and knowing: two means of access to the personal past. *Mem Cognit* **21:** 89–102.

---

[8]AC′ items were not included in this test as it would have reduced the number of usable AC trials significantly. That is, an AC′ item consists of an A-item from one AC pair and a C-item from another. To prevent carry-over effects across items, once the C-item from a pair is used to make an AC′ item, that pair cannot be tested itself. Hence, the inclusion of AC′ items would have reduced the number of items at test by half. The exclusion of AC′ items means that hit rates to AC items could be supported by either item memory (i.e., memory for just the C-items, without knowing which A-item they should correspond to) or associative memory (i.e., memory for the AC associations).

Rhodes SM, Donaldson DI. 2007. Electrophysiological evidence for the influence of unitization on the processes engaged during episodic retrieval: enhancing familiarity based remembering. *Neuropsychologia* **45:** 412–424.

Roediger HL III, Karpicke JD. 2006. The power of testing memory: basic research and implications for educational practice. *Perspect Psychol Sci* **1:** 181–210.

Rotello CM, Zeng M. 2008. Analysis of RT distributions in the remember-know paradigm. *Psychon Bull Rev* **15:** 825–832.

Rotello CM, Macmillan NA, Reeder JA, Wong M. 2005. The remember response: subject to bias, graded, and not a process-pure indicator of recollection. *Psychon Bull Rev* **12:** 865–873.

Rubin RD, Watson PD, Duff MC, Cohen NJ. 2014. The role of the hippocampus in flexible cognition and social behavior. *Front Hum Neurosci* **8:** 742.

Sadeh T, Moran R, Goshen-Gottstein Y. 2015. When items 'pop into mind': variability in temporal-context reinstatement in free-recall. *Psychon Bull Rev* **22:** 779–790.

Schacter DL, McGlynn SM. 1989. Implicit memory: effects of elaboration depend on unitization. *Am J Psychol* **102:** 151–181.

Schacter DL, Addis DR, Hassabis D, Martin VC, Spreng RN, Szpunar KK. 2012. The future of memory: remembering, imagining, and the brain. *Neuron* **76:** 677–694.

Shohamy D, Wagner AD. 2009. Integrative encoding. *Am J Psychiatry* **166:** 284.

Skinner EI, Fernandes MA. 2007. Neural correlates of recollection and familiarity: A review of neuroimaging and patient data. *Neuropsychologia* **45:** 2163–2179.

Tibon R, Vakil E, Goldstein A, Levy DA. 2012. Unitization and temporality in associative memory: evidence from modulation of context effects. *J Mem Lang* **67:** 93–105.

Tulving E. 1985. Memory and consciousness. *Can Psychol* **21:** 1–12.

Underwood BJ. 1949. Proactive inhibition as a function of time and degree of prior learning. *J Exp Psychol* **39:** 24–34.

van der Jeugd A, Goddyn H, Laeremans A, Arckens L, D'Hooge R, Verguts T. 2009. Hippocampal involvement in the acquisition of relational associations, but not in the expression of a transitive inference task in mice. *Behav Neurosci* **123:** 109–114.

Verfaellie M, Page K, Orlando F, Schacter DL. 2005. Impaired implicit memory for gist information in amnesia. *Neuropsychology* **19:** 760–769.

Wais PE, Squire LR, Wixted JT. 2010. In search of recollection and familiarity signals in the hippocampus. *J Cogn Neurosci* **22:** 109–123.

Wilson MD. 1988. The MRC psycholinguistic database: machine readable dictionary, version 2. *Behav Res Methods Instrum Comput* **20:** 6–11.

Winocur G, Kinsbourne M. 1978. Contextual cueing as an aid to Korsakoff amnesics. *Neuropsychologia* **16:** 671–682.

Winocur G, Moscovitch M. 1983. Paired-associate learning in institutionalized and noninstitutionalized old people: An analysis of interference and context effects. *J Gerontol* **38:** 455–464.

Winocur G, Weiskrantz L. 1976. An investigation of paired-associate learning in amnesic patients. *Neuropsychologia* **14:** 97–110.

Winocur G, Moscovitch M, Bruni J. 1996. Heightened interference on implicit, but not explicit, tests of negative transfer: Evidence from patients with unilateral temporal lobe lesions and normal old people. *Brain Cogn* **30:** 44–58.

Wixted JT. 2007. Signal-detection theory and the neuroscience of recognition memory. In *The foundations of remembering: essays in honor of Henry L. Roediger III* (ed. Nairne JS), pp. 67–82. Psychology Press, New York, NY.

Wixted JT, Mickes L. 2010. A continuous dual-process model of remember/know judgments. *Psychol Rev* **117:** 1025–1054.

Wixted JT, Squire LR. 2011. The medial temporal lobe and the attributes of memory. *Trends Cogn Sci* **15:** 210–217.

Wixted JT, Stretch V. 2004. In defense of the signal detection interpretation of remember/know judgments. *Psychon Bull Rev* **11:** 616–641.

Yonelinas AP. 2001. Consciousness, control, and confidence: the 3 Cs of recognition memory. *J Exp Psychol Gen* **130:** 361–379.

Yonelinas AP. 2002. The nature of recollection and familiarity: a review of 30 years of research. *J Mem Lang* **46:** 441–517.

Yonelinas AP, Jacoby LL. 1995. The relation between remembering and knowing as bases for recognition: effects of size congruency. *J Mem Lang* **34:** 622–643.

Yonelinas AP, Dobbins I, Szymanski MD, Dhaliwal HS, King L. 1996. Signal-detection, threshold, and dual-process models of recognition memory: ROCs and conscious recollection. *Conscious Cogn* **5:** 418–441.

Yonelinas AP, Kroll NE, Dobbins IG, Soltani M. 1999. Recognition memory for faces: when familiarity supports associative recognition judgments. *Psychon Bull Rev* **6:** 654–661.

Zalesak M, Heckers S. 2009. The role of the hippocampus in transitive inference. *Psychiatry Res* **172:** 24–30.

Zeithamova D, Schlichting ML, Preston AR. 2012. The hippocampus and inferential reasoning: building memories to navigate future decisions. *Front Hum Neurosci* **6:** 70.