Recognition memory and the hippocampus: A test of the hippocampal contribution to recollection and familiarity

Annette Jeneson, C. Brock Kirwan, Ramona O. Hopkins, et al.

References	This article cites 40 articles, 5 of which can be accessed free at: **http://learnmem.cshlp.org/content/17/1/63.full.html#ref-list-1**
Email alerting service	Receive free email alerts when new articles cite this article - sign up in the box at the top right corner of the article or **click here**

Recognition memory and the hippocampus: A test of the hippocampal contribution to recollection and familiarity

Annette Jeneson,[1] C. Brock Kirwan,[2] Ramona O. Hopkins,[2,3] John T. Wixted,[1] and Larry R. Squire[1,4,5,6,7]

[1]Department of Psychology, University of California, San Diego, California 92093, USA; [2]Department of Psychology and Neuroscience Center, Brigham Young University, Provo, Utah 84602, USA; [3]Department of Medicine, Pulmonary and Critical Care Division, Intermountain Medical Center, Murray, Utah 84143, USA; [4]Department of Psychiatry, University of California, San Diego, California 92093, USA; [5]Department of Neurosciences, University of California, San Diego, California 92093, USA; [6]Veterans Affairs Healthcare System, San Diego, California 92161, USA

It has been suggested that the hippocampus selectively supports recollection and that adjacent cortex in the medial temporal lobe can support familiarity. Alternatively, it has been suggested that the hippocampus supports both recollection and familiarity. We tested these suggestions by assessing the performance of patients with hippocampal lesions on recognition memory tests that differ in the extent to which recollection and familiarity contribute to the recognition decision. When targets and foils are highly similar, prior evidence suggests that, on a forced-choice test in which targets are presented together with highly similar, corresponding foils (the FC-C format), performance is supported primarily by familiarity. By contrast, when targets are presented together with foils that are similar to other targets (the FC-NC format) or when memory is tested in a yes/no (Y/N) format, performance is based much more strongly on recollection. Accordingly, a finding that hippocampal damage impaired both Y/N recognition and FC-NC recognition but spared FC-C recognition would suggest that the hippocampus selectively supports recollection. We administered Y/N, FC-C, and FC-NC tests to five memory-impaired patients with circumscribed hippocampal lesions and 14 controls. The patients were impaired on all three types of recognition test, and there was no indication that the patients were disproportionately benefited or disproportionately impaired on any test. This pattern of performance suggests that the hippocampus supports both recollection and familiarity.

[Supplemental material is available online at http://www.learnmem.org.]

The medial temporal lobe (the hippocampus plus the entorhinal, perirhinal, and parahippocampal cortices) is essential for the recognition of past experience. The capacity for recognition is widely thought to depend on two distinct processes, recollection and familiarity (Mandler 1980; Yonelinas et al. 2002; Wixted 2007). Recollection involves remembering specific details about the episode in which an item was encountered. Familiarity involves simply knowing that an item was presented, even when no additional information can be retrieved about the learning episode itself. This distinction has been prominent in recent discussions about memory, particularly in relation to its possible anatomical basis. One proposal is that recollection depends on the hippocampus and that familiarity depends on the adjacent medial temporal lobe cortex (Brown and Aggleton 2001; Mayes et al. 2002; Yonelinas et al. 2002). Alternatively, it has been suggested that the hippocampus is important for both recollection and familiarity and that the distinction between these two processes does not illuminate the functional differences between the hippocampus and adjacent cortex (Rutishauser et al. 2006; Wais et al. 2006; Squire et al. 2007; Wixted 2007).

To clarify the role of the hippocampus in recognition memory, the performance of memory-impaired patients with hippocampal damage has often been compared to that of controls on tests that differ in the extent to which recollection and familiarity contribute to the recognition memory decision. One such approach involves the use of highly similar targets and foils tested using a yes/no (Y/N) format and a forced-choice corresponding (FC-C) format (Holdstock et al. 2002; Westerberg et al. 2006; Bayley et al. 2008). In the case of the Y/N format, participants see a list of targets intermixed with foils (each similar to one of the targets) and are asked to respond "yes" to the targets and "no" to the foils. In the FC-C format, participants see a target presented together with its corresponding, similar foil and are asked to identify the target. According to an idea advanced by Norman and O'Reilly (2003), familiarity can support FC-C recognition because one can make effective use of small but consistent differences between the familiarity signals triggered by the target and foil items. That is, when targets and foils are highly similar and are presented together in the FC-C format, familiarity for the target is likely to slightly and reliably exceed that of the similar foil, allowing the target to be correctly chosen on the basis of familiarity. By contrast, in the Y/N format, slight differences between the familiarity signals of target items and their corresponding foils no longer provide useful information, because the targets and their corresponding foils are not presented together at

test. Accordingly, good performance with the Y/N format is more dependent on recollection than with the FC-C format.

The same explanation may account for why performance on FC-C tests sometimes exceeds performance on forced-choice non-corresponding (FC-NC) tests (e.g., Hintzman 1988). In the FC-NC format, target items are presented together with a noncorresponding foil that is similar to another target item from the study list (Fig. 1). Thus, the FC-C and FC-NC tests differ only in that the similar targets and foils are presented together in the former but not in the latter. When the corresponding targets and foils are presented together (in the FC-C format), participants can make effective use of consistent differences in familiarity values (as discussed above). When the corresponding targets and foils are not presented together, participants cannot make use of the small differences that they might detect through a side-by-side comparison.

A recent study (Migo et al. 2009) provides empirical support for the suggestion that familiarity primarily contributes to the recognition decision in the FC-C format, whereas recollection contributes more strongly in the Y/N format and in the FC-NC format. In the Migo et al. (2009) study, healthy participants received FC-C, FC-NC, and Y/N tests after receiving either standard instructions to make their decisions based on familiarity or recollection or modified instructions to base their decisions on familiarity only. On the FC-NC test and on the Y/N test, performance using familiarity alone was significantly worse than under standard instructions. On the FC-C test, performance using familiarity alone was nearly as good as under standard instructions. This result supports the earlier suggestion that FC-C is primarily supported by familiarity, whereas recollection plays a larger role in FC-NC and Y/N recognition (Norman and O'Reilly 2003).

If the hippocampus is critical for recollection but not for familiarity, and if good performance on the FC-C format can be achieved using familiarity alone, then patients with focal hippocampal lesions should be differentially impaired on the Y/N and FC-NC recognition test formats compared with the FC-C format. Three recent studies investigated this issue by assessing the effects of hippocampal damage (or presumed hippocampal damage) on Y/N and FC-C tests that used highly similar targets and foils (the FC-NC format was not used in these earlier studies). All three studies used black-and-white silhouette images of objects, and the FC-C test involved the target item and three similar foils (i.e., multiple-choice with four alternatives). In one study, a single patient with bilateral hippocampal damage (patient Y.R.) was impaired on the test of Y/N recognition but was unimpaired on the test of FC-C recognition (Holdstock et al. 2002). A similar

result was reported in patients with a diagnosis of mild cognitive impairment (MCI) (Westerberg et al. 2006). These findings have been taken to support the suggestion that the hippocampus selectively supports recollection. However, the study by Holdstock and colleagues (2002) involved a single patient (Y.R.), and findings from a single patient need not agree with findings obtained from a group of patients (see Discussion). In addition, the Westerberg et al. (2006) study involved individuals with a diagnosis of MCI for whom anatomical data were not available. It is therefore not clear what the status of the hippocampus was in these patients. Moreover, a standard Y/N procedure involving 12 study items and 24 test items (as opposed to the 12 study test items and 60 test items used in previous studies [Holdstock et al. 2002; Westerberg et al. 2006; Bayley et al. 2008]) might be better suited to address the question of whether Y/N recognition is selectively impaired in hippocampal patients. Bayley et al. (2008) tested five patients with circumscribed hippocampal damage using the same test materials and procedure as in Holdstock et al. (2002). When all 60 test items of the Y/N test were scored, the patients were found to be more impaired on the Y/N test than on the FC-C test. However, when only the first 24 test items of the Y/N test were scored (according to the standard Y/N procedure), the patients were found to be impaired on both the Y/N and FC-C tests to a similar degree.

The earlier studies (Holdstock et al. 2002; Westerberg et al. 2006; Bayley et al. 2008) compared Y/N recognition to four-alternative FC-C recognition in which each target was presented along with three corresponding foils. None of the earlier studies investigated how patients with hippocampal lesions performed on FC-NC recognition. As Migo et al. (2009) point out, FC-C and FC-NC tests are better matched than the FC-C and Y/N tests because they are both forced-choice tests, and both use the same number of test trials. As such, they argue, it would be more useful to compare FC-C recognition performance to FC-NC recognition performance to determine whether patients with hippocampal lesions are intact or impaired on familiarity-based tests.

The present study assessed Y/N, two-alternative FC-C, and two-alternative FC-NC recognition, using highly similar targets and foils. The tests were given to five patients with circumscribed hippocampal damage and 14 matched controls. The Y/N test involved an equal number of targets and foils (unlike the design used in the prior studies), and a two-alternative format was used for the forced-choice tests to make them comparable to the Y/N test (i.e., all three tests involved the same number of targets and foils). The FC-C and FC-NC tests were identical except with respect to how the targets and their corresponding foils were paired on the recognition test.

To extend our findings beyond the stimuli used in prior studies, we also used color photographs of objects in addition to black-and-white silhouettes. Using these three test formats, we asked the following questions: (1) Does hippocampal damage impair Y/N recognition memory but spare, or disproportionately benefit, FC-C recognition (as suggested by the view that the hippocampus supports recollection and not familiarity)? Or does hippocampal damage impair Y/N and FC-C recognition similarly (as suggested by the view that the hippocampus supports both recollection and familiarity)? (2) Does hippocampal damage disproportionately benefit FC-C recognition relative to FC-NC recognition, or does hippocampal damage impair FC-C and FC-NC recognition similarly?
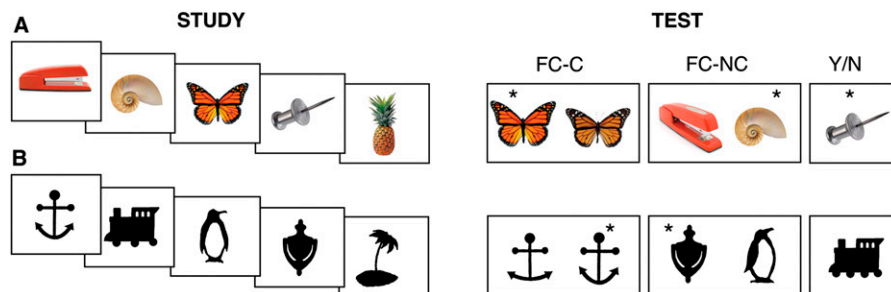

Figure 1. Test format and materials for the three kinds of recognition test. For each test, 12 images, either objects (*A*) or silhouettes (*B*), were presented at study, and memory was tested in one of three ways. For the forced-choice corresponding test (FC-C), each target (a studied item) was presented together with a highly similar foil (a new item). For the forced-choice noncorresponding test (FC-NC), each target was presented together with a foil that was highly similar to a different target from the study list. Participants were asked to point to exactly the same image they had seen during study. For the yes/no test (Y/N), the 12 targets and the 12 foils (each similar to one of the targets) were intermixed and presented one at a time. Participants were asked to respond "yes" if they had seen exactly the same image before and "no" if they had not. Asterisks identify the target items.

Results

Discriminability performance

A preliminary ANOVA, with materials (objects or silhouettes) and type of recognition test (FC-C, FC-NC, or Y/N) as repeated measures and group (memory-impaired patients or controls) as a between-subjects factor, revealed a main effect of materials ($F_{(1,34)} = 77.1$, $P < 0.001$), a main effect of test ($F_{(2,34)} = 4.66$, $P = 0.016$), and a main effect of group ($F_{(1,17)} = 24.2$, $P < 0.001$) but no significant interactions. Given the absence of an interaction with materials as a factor, and the fact that there is no theoretically important difference between the two types of material, we averaged the discriminability estimates obtained for objects and for silhouettes to obtain a single discriminability estimate for each of the three kinds of recognition test. Table 1 presents the recognition scores separately for the two types of material. The FC-C vs. FC-NC advantage is apparent for both patients and controls on both types of material. Similarly, the FC-C vs. Y/N advantage is apparent in all cases except for the controls in the objects test, where the score for the controls appeared to be suppressed by a ceiling effect.

Figure 2 shows the discriminability scores of memory-impaired patients and controls on the two types of forced-choice recognition tests (d') and the Y/N recognition test (d_a). Hit rates and false alarm rates for each condition are provided in Table 1. We first compared the performance of patients and controls on the three types of recognition tests and found an effect of group ($F_{(1,17)} = 24.01$, $P < 0.001$) and test ($F_{(2,34)} = 4.57$, $P = 0.02$) but no group × test interaction ($F_{(2,34)} = <1.0$, $P = 0.48$). The group effect reflects the fact that the patients were impaired on all three tests (although they did perform above chance in each case, all Ps < 0.01). The significant effect of test reflects the fact that patients and controls performed better on the FC-C test compared with the FC-NC test, an effect that we then explored further.

Previous work suggested that healthy participants should benefit from the FC-C format, relative to the FC-NC format (Hintzman 1988). In addition, the complementary learning systems (CLS) model of recognition memory predicts that the benefit of the FC-C format over the FC-NC format should be greater for hippocampal patients than for controls (Norman and O'Reilly 2003). Thus, we asked whether patients and controls exhibited an advantage in the FC-C format compared with the FC-NC format and whether one group benefited more than the other. An ANOVA comparing the patients and controls on the two tests revealed an effect of group ($F_{(1,17)} = 23.27$, $P < 0.001$) and an effect of test ($F_{(1,17)} = 10.30$, $P < 0.001$), but no group × test interaction ($F_{(1,17)} = 0.60$, $P = 0.45$). Thus, the FC-C procedure benefited patients and controls alike, and there was no hint that the patients were disproportionately benefited. Indeed, the patients were impaired relative to controls on the FC-C test ($t_{(4)} = 5.41$, $P = 0.006$) just as they were impaired on the FC-NC test ($t_{(4)} = 7.48$, $P = 0.002$).

We also compared performance on the Y/N and FC-C tests, based on the proposal that recollection is critical for good Y/N recognition, whereas familiarity may be sufficient for good FC-C recognition, at least when targets and foils are highly similar (O'Reilly and Rudy 2000; Mayes et al. 2002). An ANOVA comparing the patients and controls on the two tests revealed an effect of group ($F_{(1,17)} = 16.91$, $P = 0.001$) and an effect of test ($F_{(1,17)} = 5.26$, $P = 0.04$), but no group × test interaction ($F_{(1,17)} = 1.70$, $P = 0.21$). Thus, relative to the Y/N test, the FC-C test procedure benefited patients and controls similarly. Furthermore, the patients were impaired on both the Y/N and FC-C tests (Ps < 0.01).

The same results for the FC-C vs. FC-NC contrast and the FC-C vs. Y/N contrast were obtained using relative (as opposed to absolute) measures of change. We computed the ratio of each participant's FC-C performance (d') to his/her FC-NC performance (d') and to his/her Y/N performance (d_a). We then performed t-tests on both the ratio scores and the log ratio scores. In no case did the differences in these measures between patients and controls approach significance (Ps > 0.1).

Receiver operating characteristics (ROCs) analysis

We also constructed ROCs from the confidence ratings for Y/N decisions in the third testing session (see Supplemental material: ROC analysis; Fig. 2). The primary purpose for the ROC analysis was to obtain a slope parameter for each group and condition so that d_a could be used as an estimate of discriminability for Y/N recognition (see Data analysis). As described below, the slope parameter is an estimate of the ratio of the standard deviation of the foil memory-strength distribution to the standard deviation of the target memory-strength distribution. A typical value for this estimate is ~0.80 (Ratcliff et al. 1992), indicating that the variance of the target distribution exceeds the variance of the foil distribution. The slope values we obtained were considerably greater than this typical value (for silhouettes: $s = 1.21$ and 1.49; for objects: $s = 0.96$ and 1.77, for the patients and controls, respectively) (Supplemental Fig. 2). These slope values indicate that, contrary to what is typically found, the variance of the foil distribution exceeded the variance of the target distribution in our study. As described below, we suggest that this atypical result reflects a recollection process whereby foils may sometimes be incorrectly accepted as old items based on recollection of those features of the foil that match the target (i.e., foils are endorsed as targets based on a recall-to-accept process) while overlooking diagnostic features that distinguish the foil from its highly similar target.

Discussion

We investigated the role of the hippocampus in recollection and familiarity. One proposal is that the hippocampus selectively

Table 1. Mean hit and false alarm (FA) rates and discriminability scores for three test formats and two types of material

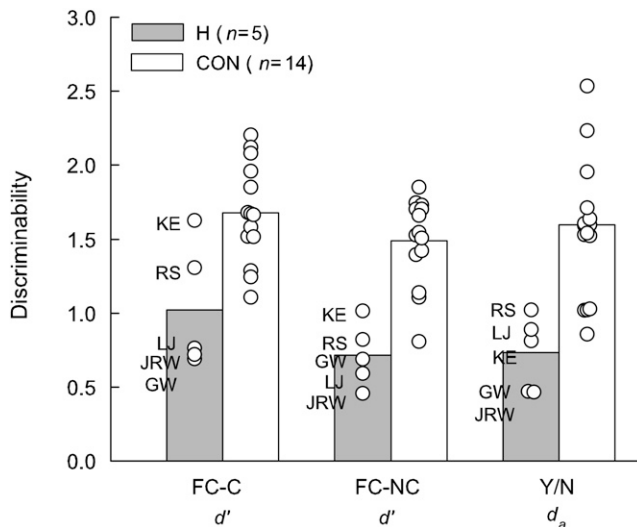|	Patients			Controls		
	FC-C (d')	FC-NC (d')	Y/N (d_a)	FC-C (d')	FC-NC (d')	Y/N (d_a)
Objects						
d' or d_a	1.40 (0.21)	0.89 (0.17)	1.03 (0.17)	2.07 (0.12)	1.93 (0.13)	2.14 (0.18)
Hit rate	0.83 (0.03)	0.73 (0.04)	0.87 (0.02)	0.92 (0.01)	0.90 (0.02)	0.94 (0.02)
FA rate	0.17 (0.03)	0.27 (0.04)	0.55 (0.03)	0.08 (0.01)	0.10 (0.02)	0.24 (0.03)
Silhouettes						
d' or d_a	0.60 (0.19)	0.52 (0.09)	0.42 (0.18)	1.22 (0.11)	0.97 (0.06)	1.07 (0.11)
Hit rate	0.66 (0.05)	0.64 (0.02)	0.80 (0.06)	0.80 (0.02)	0.75 (0.02)	0.85 (0.02)
FA rate	0.34 (0.05)	0.36 (0.02)	0.66 (0.03)	0.20 (0.02)	0.25 (0.02)	0.43 (0.03)

Parentheses show SEM.

Figure 2. Performance by patients with circumscribed hippocampal lesions (H) and controls (CON) on the three kinds of recognition memory tests: forced-choice corresponding (FC-C), forced-choice noncorresponding (FC-NC), and yes/no (Y/N). Discriminability was measured by d' for the two-forced choice tests, and by d_a for the Y/N test (see Data analysis). Each participant's score is represented by a circle, and patients are identified by initials. The scores for each type of recognition test are the means of eight different tests, four constructed from objects and four constructed from silhouettes.

supports recollection (Brown and Aggleton 2001; Mayes et al. 2002; Yonelinas et al. 2002). An alternative proposal is that the hippocampus supports both recollection and familiarity (Rutishauser et al. 2006; Squire et al. 2007). To test these proposals, we assessed the performance of five memory-impaired patients with circumscribed hippocampal lesions on recognition tests that differ in the extent to which they depend on recollection and familiarity. Three recognition tests were used (Y/N, FC-C, and FC-NC), and all three involved highly similar targets and foils (Fig. 1). Good performance on the FC-C test can be achieved by relying primarily on familiarity, whereas good performance on both the FC-NC and the Y/N tests depends on recollection (Norman and O'Reilly 2003; Migo et al. 2009). Accordingly, if the hippocampus selectively supports recollection, then patients with hippocampal damage should be disproportionately benefited by tests of FC-C recognition compared with tests of FC-NC or Y/N recognition. Our results ran counter to this prediction. The patients were impaired on all three kinds of recognition test, and the extent of impairment was similar across the three tests.

The idea that familiarity primarily contributes to recognition memory decisions in the FC-C format, whereas recollection contributes more to the recognition memory decision in FC-NC and Y/N formats, was first advanced by Norman and O'Reilly (2003) as part of their influential CLS model of recognition memory. Recently, Migo et al. (2009) provided direct evidence in support of that idea by showing that participants instructed to respond on the basis of familiarity alone were much less impaired in the FC-C format compared with the Y/N and FC-NC formats. If performance in the FC-C format is in fact based primarily on familiarity, and if performance in the FC-NC and Y/N formats is based more on recollection, then our finding that patients were similarly impaired across all three tests suggests that the hippocampus supports the familiarity component of recognition memory as well as the recollection component (also see Rutishauser et al. 2006, 2008; Wais et al. 2006). This finding is also consistent with an earlier finding suggesting that the four patients in the

current study (J.R.W., G.W., L.J, and R.S.) who were tested on the Doors and People test were equally impaired on the recall and recognition portions (Manns et al. 2003).

In contrast to the findings from the current study and our earlier study (Bayley et al. 2008), two earlier studies (Holdstock et al. 2002; Westerberg et al. 2006) reported impaired Y/N recognition and unimpaired FC-C recognition in their patients. Two points deserve comment. First, neuroanatomical data were not available for the patients with MCI examined by Westerberg et al. (2006). In this diagnostic group, it is difficult to make inferences about the contribution of the hippocampus to test performance in the absence of anatomical information. Second, the other study (Holdstock et al. 2002) involved a single patient (Y.R.). Findings from a single patient need not agree with findings obtained from a larger group of patients. Note that patient K.E. in our study performed similarly to patient Y.R. (Fig. 2). Specifically, both K.E. and Y.R. performed considerably better on the FC-C test than on the Y/N test (patient K.E., Y/N d_a = 0.82; Y/N d' = 0.86; FC d' = 1.63; patient Y.R., Y/N d' = 0.94; FC d' = 1.55). It is difficult to place a strong interpretation on K.E.'s performance here, because he did not perform better on the FC-C test than on the Y/N test in our earlier study (Bayley et al. 2008).

Alternatively, it is possible that this kind of variation is evident in unimpaired individuals as well as in patients. Note, for example, that half of the controls in our experiment performed numerically better on the FC-C test compared with both the Y/N test and the FC-NC test. Furthermore, for two of these controls, the FC-C advantage was considerable (similar to the advantage observed for patients K.E. and Y.R.). The source of these individual differences in test performance is not known, but it is important to note that the differences between FC-C performance and FC-NC or Y/N performance for these two individuals was obtained even in the absence of hippocampal damage. Thus, the performance of patients Y.R. and K.E. could reflect individual differences in test performance (which would have been present premorbidly), rather than an effect of a particular kind of hippocampal pathology. Similarly, if a patient had an advantage for recollection-based responding premorbidly, then a selective deficit in recollection might give the impression that familiarity and recollection are equally impaired after hippocampal damage. These considerations emphasize the value of group studies for obtaining information about the typical effects of hippocampal damage.

The methods employed in the present study provided improved conditions for asking whether patients with hippocampal damage as a group would be differentially impaired across tests that differ in the extent to which they rely on recollection and familiarity. First, the hippocampal lesions of the patients in the present study were rigorously measured using quantitative volumetric analysis of magnetic resonance images (see Materials and Methods) (Bayley et al. 2005; Gold and Squire 2005). Second, we included a FC-NC condition in addition to the Y/N and FC-C conditions that had been used in previous studies. The FC-C and FC-NC tests are better matched than the FC-C and Y/N tests because they are both forced-choice tests and use the same number of test trials (a point also emphasized by Migo et al. 2009). Third, we obtained confidence ratings so that Y/N discriminability could be measured using d_a, which allows for the most appropriate comparison between Y/N performance and FC-C and FC-NC performance. Fourth, we asked whether the results obtained using this procedure would generalize to a stimulus set other than the one used in all three prior studies investigating this issue. We found that the patients were similarly impaired across all three recognition tests. We therefore conclude that the hippocampus supports both recollection and familiarity.

Other methods have also been used to assess recollection and familiarity in hippocampal patients or patients with damage to

related structures (e.g., recall vs. recognition, remember vs. know, ROC analysis, structural equation modeling, the process-dissociation procedure, item memory vs. source memory, event-related potential indices of recollection), and the results have been variable and often depend on single cases (Düzel et al. 2001; Yonelinas et al. 2002; Holdstock et al. 2002, 2005; Mayes et al. 2002, 2004; Manns et al. 2003; Bastin et al. 2004; Quamme et al. 2004; Aggleton et al. 2005; Barbeau et al. 2005; Cipolotti et al. 2006; Wais et al. 2006; Tsivilis et al. 2008; Turriziani et al. 2008; Vann et al. 2008; Verfaellie et al. 2008). Although some of these studies have presented evidence suggesting that the hippocampus selectively subserves recollection, many of them used controversial methods that may confound the assessment of recollection and familiarity with the assessment of strong and weak memories (see Wixted 2009). This difficulty is of particular concern for studies that rely on model-based methods, such as the remember/know procedure and ROC analysis (i.e., methods that depend on specific quantitative models that assume that recollection is a threshold process, not a continuous process) (Mickes et al. 2009). Thus, instead of showing that hippocampal lesions selectively impair recollection, the results obtained with these methods may instead show that hippocampal lesions reduce memory strength, whether memory is based on recollection or on familiarity (Squire et al. 2007; Wixted 2007). Excluding model-based studies (because of the possible strength confound) and focusing on studies that used a group design, a somewhat more consistent picture emerges about the effects of hippocampal damage. For example, group studies comparing recall and recognition have found that performance on both tasks is similarly impaired by hippocampal lesions (for a review, see Squire et al. 2007). Because it is widely assumed that recall is supported only by recollection and that recognition can be supported by recollection and familiarity, comparable deficits on these two tasks suggest that the hippocampus subserves both recollection and familiarity (for evidence that mammillary body atrophy impairs recall more than recognition, see Tsivilis et al. 2008).

None of the prior studies using FC and Y/N recognition tests to investigate the role of the hippocampus in recollection and familiarity found a performance advantage among the healthy controls for the FC-C format compared with the Y/N format (Holdstock et al. 2002; Westerberg et al. 2006; Bayley et al. 2008), as we did in our study. Unlike the comparison between the FC-C and FC-NC formats (which was not made in the earlier studies), a direct comparison between the FC-C and Y/N formats is complicated by measurement issues. Typically, this comparison is made using d'. However, the Y/N d' table assumes an equal-variance signal-detection model, which is to say that it assumes that the ROC (were it assessed) would be symmetrical. When this assumption is not correct, then the Y/N and forced-choice d' values are not directly comparable, and it is not possible to determine whether performance in the FC-C condition exceeds that in the Y/N condition. In the current study, ROC data were collected, the ROC was not symmetrical (as the d' measure assumes), and Y/N recognition was therefore measured by d_a, which is the more appropriate measure in this case (Macmillan and Creelman 2005). When d_a was used to measure Y/N recognition, performance was approximately the same in the Y/N and FC-NC conditions for both patients and controls. (This is the expected result in light of previous research showing that Y/N and two-alternative FC performance is comparable once the asymmetry of the Y/N ROC is taken into consideration) (Jang et al. 2009). In addition, the expected performance advantage of the FC-C over the Y/N condition was now evident for both patients and controls. Even so, hippocampal lesions impaired Y/N and FC-C recognition memory to the same extent, and the FC-C condition did not differentially benefit the patients.

The ROC analyses of our Y/N data, in addition to permitting the computation of d_a, yielded an interesting and unexpected result. Specifically, the Y/N recognition slope estimates obtained in the present study were greater than 1 instead of being near 0.80 as is typically found. We suggest that slope estimates greater than 1 might be peculiar to Y/N test formats that involve highly similar targets and foils. To our knowledge, only one other study has presented a Y/N ROC curve having a slope similar to the Y/N ROC curve obtained in the present study (i.e., Rotello et al. 2000). In that study, common nouns and their highly similar plurals (e.g., *chair* and *chairs*) were used as targets and foils on the recognition test. Although ROC slopes were not reported, we estimated from their data that the slope of their ROC curve was also greater than 1 (~1.12). Slope estimates greater than 1 derived from tests involving high target-foil similarity may reflect the fact that, when targets and foils are highly similar, foils can be endorsed as old based on a recall-to-accept process. That is, foils may sometimes be incorrectly accepted as old when a subject correctly recollects features of the foil that match the target but overlooks ways in which the target and foil differ. A recall-to-accept process, peculiar to tasks involving highly similar targets and foils, was also identified by Migo et al. (2009). Using a remember/know procedure in which participants were asked how they made the Y/N recognition judgment after each trial, Migo et al. (2009) found that 65.0% of the false alarms (i.e., 65.0% of the trials in which a new item was mistakenly recognized as old) were based on a recall-to-accept process for foils. This concept of a recall-to-accept process for foils, and the way in which it accounts for an ROC slope greater than 1, is discussed further in the Supplemental material: ROC analysis.

In summary, the question of interest was whether the hippocampus is critical for recollection but not for familiarity, or whether the hippocampus is critical for both recollection and familiarity. To address this question, we gave three kinds of recognition memory test (FC-C, FC-NC, and Y/N) to patients with focal hippocampal lesions and controls. Recent evidence (Migo et al. 2009) and our own finding of an FC-C performance advantage (which has not been previously documented) supports the original suggestion by Norman and O'Reilly (2003) that these tests differ in the extent to which they depend on familiarity or recollection. Specifically, familiarity primarily contributes to recognition memory decisions in the FC-C format, whereas recollection contributes more strongly to recognition memory decisions in the FC-NC format and the Y/N format. If the hippocampus is critical for recollection but not for familiarity, then patients with hippocampal damage should be impaired on Y/N and FC-NC recognition memory tests but should perform noticeably better or even be intact on FC-C recognition tests. However, if the hippocampus is critical for both recollection and familiarity, then patients with hippocampal damage should be impaired similarly across all three tests. Because the patients were similarly impaired across all three recognition tests, we conclude that the hippocampus is involved in both of the component processes that support the ability to recognize past experience, namely, recollection and familiarity.

Materials and Methods

Participants

Five memory-impaired patients participated (Table 2). The same patients participated in our previous study (Bayley et al. 2008). All five patients have bilateral lesions thought to be limited to the hippocampus (CA fields, dentate gyrus, plus subicular complex). K.E. has been diagnosed with diabetes (type II), but all patients are otherwise in good general health. K.E. became amnesic in 2004

Table 2. Characteristics of memory-impaired patients

					WMS-R				
Patient	Gender	Age (yr)	Education (yr)	IQ	Attention	Verbal	Visual	General	Delay
K.E.	M	67	13.5	108	114	64	84	72	55
L.J.	F	71	12	101	105	83	60	69	<50
G.W.	M	49	12	108	105	67	86	70	<50
J.R.W.	M	45	12	90	87	65	95	70	<50
R.S.	M	52	12	99	99	85	81	82	<50

IQs for patients K.E., L.J., and G.W. are WAIS-III; IQs for patients J.R.W. and R.S. are WAIS-R.

after an episode of ischemia associated with kidney failure and toxic shock syndrome. L.J. (the only female) became amnesic in 1988 during a 6-mo period with no known precipitating event. Her memory impairment has been stable since that time. Patients G.W. and R.S. became amnesic in 2001 and 1998, respectively, following a drug overdose and associated respiratory failure. J.R.W. became amnesic in 1990 following an anoxic episode associated with cardiac arrest. Immediate and delayed (12 min) recall of a short prose passage averaged 3.8 and 0.4 segments (21 segments maximum) (Gilbert et al. 1968).

Estimates of medial temporal lobe damage were based on quantitative analysis of magnetic resonance images from 19 healthy males for the four male patients and 11 females for patient L.J. (Gold and Squire 2005). K.E., L.J., R.S., G.W., and J.R.W. have an average bilateral reduction in hippocampal volume of 49%, 46%, 33%, 48%, and 44%, respectively (all values >3 SD from the control mean). Each patient has similar reductions on the left and right sides (always within 10% of each other; mean, 7%). On the basis of two patients (L.M. and W.H.) with similar bilateral volume loss in the hippocampus for whom detailed post-mortem neuro-histologic information was obtained (Rempel-Clower et al. 1996), this degree of volume loss likely reflects nearly complete loss of hippocampal neurons. The volume of the parahippocampal gyrus, by contrast, is reduced by 17%, −8%, 1%, 12%, and 6%, respectively (all values within 2 SD of the control mean).

Additional measurements, based on four controls for each patient, were carried out for the frontal lobes, lateral temporal lobes, parietal lobes, occipital lobes, insular cortex, and fusiform gyrus (Bayley et al. 2005, 2006). The only volume reduction in these regions >1.3 SD of the control mean was the parietal lobe of patient R.S. This likely reflects natural variation in the parietal lobe volume (see Bayley et al. 2006). Nine coronal magnetic resonance images from each patient, together with detailed descriptions of the lesions, appear in Supplemental material Figure 1.

Fourteen controls also participated (11 males; mean age, 61.2 yr; mean education, 14.4 yr). For the controls, immediate and delayed prose recall averaged 8.1 and 7.0 segments, respectively.

Materials and procedure

Recognition memory tests were constructed from two types of material: color photographs of objects and black-and-white silhouette images of objects (hereafter referred to as objects and silhouettes, respectively) (Fig. 1). The objects consisted of 144 pairs of color photographs on a white background, each of which was paired with a highly similar image (total, 288 images). The similarity of the object target and foil pairs were not formally assessed. Half of the photographs depicted objects typically found outdoors, and half depicted objects typically found indoors. The silhouettes consisted of 144 pairs of images, half of which depicted natural objects and half of which depicted man-made objects. Forty-eight of the silhouette pairs were constructed from the set used by Bayley et al. (2008) (the two most similar images of the four versions of each object were selected from this set). Ninety-six additional pairs were constructed by first passing images from the Internet through a threshold filter in MATLAB to obtain black-and-white images and then modifying these images with Adobe Photoshop.

The study phase was identical for all the tests that were given in each of the three recognition formats (FC-C, FC-NC, and Y/N). Participants studied 12 images presented one at a time for 3 sec in the center of a computer screen (500-msec intertrial interval). The image on the screen subtended ~2°, and participants were seated ~45 cm from the screen. While the image was on the screen, participants made an indoor/outdoor or natural/man-made judgment for each object or silhouette. After all 12 images had been presented, each image was presented a second time for 3 sec each. This time participants were encouraged to "study the details of each picture carefully" in anticipation of the subsequent memory test. The instructions emphasized the importance of studying the details of each picture, because the recognition test would require discrimination between studied items and items that were very similar to the studied items. The recognition tests were always given 45 sec after study. There was no time constraint for responding.

For the FC-C test, participants viewed two items presented next to each other. Each display included the target item and its corresponding highly similar foil (Fig. 1). For the FC-NC test, each display included a target item and a foil that was highly similar to a different target item presented during the study phase. Participants were asked to point to exactly the same image they had seen during study. For the Y/N test, the 12 targets and the 12 foils (each similar to one of the targets) were intermixed and presented one at a time. Participants were asked to respond "yes" if they had seen exactly the same image before and "no" if they had not. In the third test session, instead of using the standard Y/N procedure, we assessed Y/N recognition memory using confidence ratings in order to construct ROC curves and obtain additional information about performance (see next section).

For each of three kinds of recognition test (FC-C, FC-NC, and Y/N), participants completed four separate tests with objects and four with silhouettes (a total of 12 tests with objects and 12 with silhouettes). These tests were distributed across three test sessions, separated by at least 3 h (patient mean, 22.8 d; control mean, 33.7 d). Six recognition tests were given in the first session (the FC-C, FC-NC, and Y/N tests were each given twice with either objects or silhouettes). Six recognition tests were also given in the second test session, but using whichever stimulus type (objects or silhouettes) was not used in the first session. In the third session, 12 recognition tests were given (the FC-C, FC-NC, and Y/N tests were each given twice with objects and twice with silhouettes).

The order in which the three recognition tests and the two kinds of stimuli (objects or silhouettes) were used was counterbalanced across participants. The order ensured that each type of recognition test was equally often given first, second, and third, and that each image pair was equally likely to be used for each type of recognition test. Which stimuli were presented as targets and which were presented at foils, and the order in which targets and foils were presented at test, was randomized for each subject. The order of presentation of items at study and test was also randomized. Practice of the study-test format using six novel silhouettes and six novel objects was given immediately before the study phase of each recognition test. Practice sessions followed the same formats as the FC-C, FC-NC, and Y/N study-test formats, except that half as many items were given for study and test.

Data analysis

Results from the three recognition tests were analyzed using the discriminability estimate d' for forced-choice recognition, and d_a for Y/N recognition. Following the method of Macmillan and Creelman (2005), FC-C and FC-NC d' values were derived for each participant by dividing the uncorrected d' score [z(hit rate) − z(false alarm rate)] by $\sqrt{2}$. d' is an appropriate calculation for recognition memory if an equal-variance signal detection (EVSD)

model provides a valid account of the data. The EVSD model invariably provides a valid account of forced-choice data. However, for Y/N recognition, an unequal-variance signal detection (UVSD) (Egan 1958) model applies. In this case, d_a is a better estimate of discriminability than d' (Macmillan and Creelman 2005). In our study, the effect of using d_a was to yield a slightly lower score than d'. Whereas d' is an estimate of the distance between the mean of the target and foil distributions divided by the common standard deviation of both distributions (the standard deviations are assumed to be equal according to the EVSD model), d_a is an estimate of the distance between the mean of the target and foil distributions divided by the root mean square of the variance of the two distributions (i.e., it takes into account the fact that these two distributions have unequal variances). d_a was calculated using the following equation (Macmillan and Creelman 2005):

$$d_a = \left(\frac{2}{1+s^2}\right)^{1/2} [z(H) - sz(F)],$$

where s is the ratio of the standard deviation of the foil distribution to the standard deviation of the target distribution (i.e., the slope parameter), H is the hit rate, and F is the false alarm rate.

In order to obtain the slope parameter we constructed ROCs from the confidence ratings for Y/N decisions in the third testing session. The ROC data were analyzed using the longstanding UVSD (see Supplemental material: ROC analysis; Egan 1958). The slope parameter was obtained by conducting maximum likelihood estimation (Ogilvie and Creelman 1968) of each group's ROC data obtained from the two Y/N recognition tests for objects and the two for silhouettes that were administered in the third test session (i.e., half of all the Y/N recognition tests). The slope parameter was estimated for each group of participants and for each type of stimulus material (objects and silhouettes). For the Y/N recognition tests, participants used a six-point confidence scale to indicate whether they recognized an item as new or old (1, definitely new; 2, probably new; 3, maybe new; 4, maybe old; 5, probably old; 6, definitely old).

A standard correction was used when the hit rate was 1.0 or the false alarm rate was 0 for any single type of recognition test in any one testing session. We converted hit rates of 1.0 to $1 - 1/(2N)$, and we converted false alarm rates of 0 to $1/(2N)$, where N is the number of targets or the number of foils, respectively. Corrections were rarely used to estimate d' for tests constructed from silhouettes, but were used more often to estimate d' for tests constructed from objects. For tests constructed from objects, corrections were used for 24.0% and 3.6% of the d' estimates for controls and patients, respectively.

Acknowledgments

References

Aggleton JP, Vann SD, Denby C, Dix S, Mayes AR, Roberts N, Yonelinas AP. 2005. Sparing of the familiarity component of recognition memory in a patient with hippocampal pathology. *Neuropsychologia* **43:** 1810–1823.

Barbeau EJ, Felician O, Joubert S, Sontheimer A, Ceccaldi M, Poncet M. 2005. Preserved visual recognition memory in an amnesic patient with hippocampal lesions. *Hippocampus* **15:** 587–596.

Bastin C, Van der Linden M, Charnallet A, Denby C, Montaldi D, Roberts N, Mayes AR. 2004. Dissociation between recall and recognition memory performance in an amnesic patient with hippocampal damage following carbon monoxide poisoning. *Neurocase* **10:** 330–344.

Bayley PJ, Gold JJ, Hopkins RO, Squire LR. 2005. The neuroanatomy of remote memory. *Neuron* **46:** 799–810.

Bayley PJ, Hopkins RO, Squire LR. 2006. The fate of old memories after medial temporal lobe damage. *J Neurosci* **26:** 13311–13317.

Bayley PJ, Wixted JT, Hopkins RO, Squire LR. 2008. Yes/no recognition, forced-choice recognition, and the human hippocampus. *J Cogn Neurosci* **20:** 505–512.

Brown MW, Aggleton JP. 2001. Recognition memory: What are the roles of the perirhinal cortex and hippocampus? *Nat Rev Neurosci* **2:** 51–61.

Cipolotti L, Bird C, Good T, Macmanus D, Rudge P, Shallice T. 2006. Recollection and familiarity in dense hippocampal amnesia: A case study. *Neuropsychologia* **44:** 489–506.

Düzel E, Vargha-Khadem F, Heinze HJ, Mishkin M. 2001. Brain activity evidence for recognition without recollection after early hippocampal damage. *Proc Natl Acad Sci* **98:** 8101–8106.

Egan JP. 1958. *Recognition memory and the operating characteristic* (Tech. note AFCRC-TN-58-5). Hearing and Communication Laboratory, Indiana University, Bloomington, IN.

Gilbert J, Levee R, Catalano K. 1968. A preliminary report on a new memory scale. *Percept Mot Skills* **27:** 277–278.

Gold JJ, Squire LR. 2005. Quantifying medial temporal lobe damage in memory-impaired patients. *Hippocampus* **15:** 79–85.

Hintzman DL. 1988. Judgments of frequency and recognition memory in multiple-trace memory model. *Psychol Rev* **95:** 528–551.

Holdstock JS, Mayes AR, Roberts N, Cezayirli E, Isaac CL, O'Reilly RC, Norman K. 2002. Under what conditions is recognition spared relative to recall after selective hippocampal damage in humans? *Hippocampus* **12:** 341–351.

Holdstock JS, Mayes AR, Gong Q, 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.

Jang Y, Wixted J, Huber DE. 2009. Testing signal-detection models of yes/no and two-alternative forced-choice recognition memory. *J Exp Psychol Gen* **138:** 291–306.

Macmillan NA, Creelman CD. 2005. *Detection theory: A user's guide*. Lawrence Erlbaum Associates, Mahwah, NJ.

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

Manns JR, Hopkins RO, Reed JM, Kitchener EG, Squire LR. 2003. Recognition memory and the human hippocampus. *Neuron* **37:** 171–180.

Mayes AR, Holdstock JS, Isaac CL, Hunkin NM, Roberts N. 2002. Relative sparing of item recognition memory in a patient with adult-onset damage limited to the hippocampus. *Hippocampus* **12:** 325–340.

Mayes AR, Holdstock JS, Isaac CL, Montaldi D, Grigor J, Gummer A, Cariga P, Downes JJ, Tsivilis D, Gaffan D, et al. 2004. Associative recognition in a patient with selective hippocampal lesions and relatively normal item recognition. *Hippocampus* **14:** 763–784.

Mickes L, Wais PE, Wixted JT. 2009. Recollection is a continuous process: Implications for dual process theories of recognition memory. *Psychol Sci* **20:** 509–515.

Migo E, Montaldi D, Norman KA, Quamme J, Mayes A. 2009. The contribution of familiarity to recognition memory is a function of test format when using similar foils. *Q J Exp Psychol* **62:** 1198–1215.

Norman KA, O'Reilly RC. 2003. Modeling hippocampal and neocortical contributions to recognition memory: A complementary-learning-systems approach. *Psychol Rev* **110:** 611–646.

Ogilvie JC, Creelman CD. 1968. Maximum-likelihood estimation of receiver-operating characteristic curve parameters. *J Math Psychol* **5:** 377–391.

O'Reilly RC, Rudy JW. 2000. Computational principles of learning in the neocortex and hippocampus. *Hippocampus* **10:** 389–397.

Quamme JR, Yonelinas AP, Kroll NEA, Sauve MJ, Knight RT. 2004. Recall and recognition in mild hypoxia: Using covariance structural modeling to test competing theories of explicit memory. *Neuropsychologia* **42:** 672–691.

Ratcliff R, Sheu CF, Gronlund SD. 1992. Testing global memory models using ROC curves. *Psychol Rev* **99:** 518–535.

Rempel-Clower NL, Zola SM, Squire LR, Amaral DG. 1996. Three cases of enduring memory impairment after bilateral damage limited to the hippocampal formation. *J Neurosci* **16:** 5233–5255.

Rotello CM, Macmillan NA, Van tassel G. 2000. Recall-to-reject in recognition: Evidence from ROC curves. *J Mem Lang* **34:** 67–88.

Rutishauser U, Mamelak AN, Schuman EM. 2006. Single-trial learning of novel stimuli by individual neurons of the human hippocampus-amygdala complex. *Neuron* **49:** 805–813.

Rutishauser U, Schuman EM, Mamelak AN. 2008. Activity of human hippocampal and amygdala neurons during retrieval of declarative memories. *Proc Natl Acad Sci* **105:** 329–334.

Squire LR, Wixted JT, Clark RE. 2007. Recognition memory and the medial temporal lobe: A new perspective. *Nat Rev Neurosci* **8:** 872–883.

Tsivilis D, Vann SD, Denby C, Roberts N, Mayes AR, Montaldi D, Aggleton JP. 2008. A disproportionate role for the fornix and mammillary bodies in recall versus recognition. *Nat Neurosci* **5:** 834–842.

Turriziani P, Serra L, Fadda L, Caltagirone C, Carlesimo GA. 2008. Recollection and familiarity in hippocampal amnesia. *Hippocampus* **18:** 469–480.

Vann SD, Tsivilis D, Denby C, Quamme JR, Yonelinas AP, Aggleton JP, Montaldi D, Mayes AR. 2008. Impaired recollection but spared familiarity in patients with extended hippocampal system damage revealed by 3 convergent methods. *Proc Natl Acad Sci* **106:** 5442–5447.

Verfaellie M, Rajaram S, Fossum K, Williams L. 2008. Not all repetition is alike: Different benefits of repetition in amnesia and normal memory. *J Int Neuropsychol Soc* **14:** 365–372.

Wais PE, Wixted JT, Hopkins RO, Squire LR. 2006. The hippocampus supports both the recollection and the familiarity components of recognition memory. *Neuron* **49:** 459–466.

Westerberg CE, Paller KA, Weintraub S, Mesulam MM, Holdstock JS, Mayes AR, Reber P. 2006. When memory does not fail: Familiarity-based recognition in mild cognitive impairment and Alzheimer's disease. *Neuropsychology* **20:** 193–205.

Wixted JT. 2007. Dual-process theory and signal-detection theory of recognition memory. *Psychol Rev* **114:** 152–176.

Wixted JT. 2009. Remember/know judgments in cognitive neuroscience: An illustration of the underrepresented point of view. *Learn Mem* **16:** 406–412.

Yonelinas AP, Kroll NE, Quamme JR, Lazzara MM, Sauve MJ, Widaman KF, Knight RT. 2002. Effects of extensive temporal lobe damage or mild hypoxia on recollection and familiarity. *Nat Neurosci* **5:** 1236–1241.