Yes/No Recognition, Forced-choice Recognition, and the Human Hippocampus

P. J. Bayley[1], J. T. Wixted[1], R. O. Hopkins[2,3], and L. R. Squire[1,4]

Abstract

■ Two recent studies reported that yes/no recognition can be more impaired by hippocampal lesions than forced-choice recognition when the targets and foils are highly similar. This finding has been taken in support of two fundamental proposals: (1) yes/no recognition tests depend more on recollection than do forced-choice tests; and (2) the hippocampus selectively supports the recollection process. Using the same stimulus materials as in the earlier studies, we tested five memory-impaired patients with circumscribed hippocampal lesions and 15 controls. As in the earlier studies, participants studied 12 pictures of objects and then took either a 12-item forced-choice test with four alternatives or a 60-item yes/no test. Patients were impaired on both tests but did more poorly on the yes/no test. However, a yes/no test based on 12 study items would conventionally involve only 24 test items (i.e., 12 study items and 12 foil items). When we scored only the first 24 test items, the patients performed identically on the yes/no and forced-choice tests. Examination of the data in blocks of 12 trials indicated that the scores of the patients declined as testing continued. We suggest that a yes/no test of 60 items is difficult relative to a 12-item forced-choice test due to the increased study–test delay and due to increased interference, not because of any fundamental difference between the yes/no and forced-choice formats. We conclude that hippocampal lesions impair yes/no and forced-choice recognition to the same extent. ■

INTRODUCTION

Recognition memory refers to the capacity to judge an item as having been encountered previously. Recognition memory is typically assessed by either a forced-choice or a yes/no procedure. In the yes/no procedure, participants are shown target items and foils one at a time and are asked to respond ''yes'' to the targets and ''no'' to the foils. In the forced-choice procedure, participants are shown a target and a foil together and are asked to identify the target. When the same study items are followed by separate forced-choice and yes/no tests, percent correct scores typically indicate an advantage for the forced-choice procedure (Kroll, Yonelinas, Dobbins, & Frederick, 2002; MacMillan & Creelman, 1991). However, when performance is expressed as d' (a measure of the ability to discriminate targets from foils), the yes/no and forced-choice procedures yield equivalent results (Smith & Duncan, 2004; Kroll et al., 2002; Khoe, Kroll, Yonelinas, Dobbins, & Knight, 2000; Yonelinas, Hockley, & Murdock, 1992; Green & Moses, 1966).

Recognition memory is thought to consist of two component processes, recollection and familiarity (Wixted, 2007; Yonelinas et al., 2002; Mandler, 1980). Recollection involves remembering specific details about the episode in which an item was encountered, and familiarity involves simply knowing that an item was presented, even when no contextual information can be retrieved.

One interesting proposal is that the yes/no and forced-choice procedures may assess recollection and familiarity differently (Bastin & Van der Linden, 2003; Aggleton & Shaw, 1996; Parkin, Yeomans, & Bindschaedler, 1994). By this view, individuals can discriminate a target from a foil on a forced-choice test on the basis of relative familiarity. In contrast, on a yes/no test, successful performance involves some degree of recollection. This view has been tested in patients with memory impairment due to hippocampal damage because of the proposal (albeit controversial) that the hippocampus selectively supports recollection (Yonelinas et al., 2002; Brown & Aggleton, 2001; for an alternative interpretation, see Squire, Wixted, & Clark, 2007; Rutishauser, Mamelak, & Schuman, 2006; Wais, Wixted, Hopkins, & Squire, 2006; Wixted & Squire, 2004). Accordingly, if the yes/no recognition test depends more on recollection than the forced-choice test does, and if patients with hippocampal damage have impaired recollection, then the patients should be more impaired on the yes/no test than on the forced-choice test. One study examined patients with left hippocampal damage as the result of stroke (Khoe et al., 2000) and found them to

[1]University of California, San Diego, [2]Brigham Young University, Provo, UT, [3]LDS Hospital, Salt Lake City, UT, [4]VA Healthcare System, San Diego, CA

be equally impaired on forced-choice and yes/no tests. The same finding was subsequently obtained in a single-case study (patient Y.R.) (Mayes, Holdstock, Isaac, Hunkin, & Roberts, 2002). Thus, the available data from patients with hippocampal damage taking conventional recognition tests do not reveal a difference in forced-choice and yes/no performance.

Another proposal has been that differences between forced-choice and yes/no performance might emerge when the targets and foils are very similar (Mayes et al., 2002; O'Reilly & Rudy, 2000). The idea in this case is that the foils produce a strong familiarity signal. In the yes/no format, the strong familiarity of the foils will result in a large number of false alarms that can be overcome only by engaging in recollection. In the forced-choice format, each target item will typically have a familiarity signal that is slightly, but reliably, stronger than a similar foil. Under these conditions, one might expect that memory-impaired patients with hippocampal damage will do better on the forced-choice test, where familiarity can be used, than on the yes/no test where recollection is important (O'Reilly & Rudy, 2000).

These ideas have been tested with a specially constructed picture recognition test involving highly similar targets and foils (3 foils for each target). Two studies have used this test and the same stimulus materials to examine the performance of memory-impaired patients (Westerberg et al., 2006; Holdstock et al., 2002). The results were that the patients were impaired on the yes/no test but were intact on the forced-choice test. The two studies attributed these findings to hippocampal dysfunction or to pathology in the hippocampus and entorhinal cortex (Westerberg et al., 2006 acknowledged the lack of neuroanatomical data as a limitation of their study). In any case, the results were not conclusive. First, one study (Holdstock et al., 2002) involved a single patient (Y.R.), and the other study (Westerberg et al., 2006) involved eight individuals with a diagnosis of mild cognitive impairment (MCI) for whom there were no anatomical data. Second, both studies employed an unconventional version of the yes/no test. Instead of n study items followed by $2n$ test items, the yes/no test in these two studies consisted of n study items and $5n$ test items ($n = 12$). The forced-choice test, on the other hand, was conventional in the sense that 12 study items were followed by 12 test items, where each test item included a target and three foils. As a result of this difference in the number of test items, the yes/no test was substantially longer than the forced-choice test and also could be expected to generate interference. Either of these factors might have disadvantaged memory-impaired patients.

We tested five memory-impaired patients with circumscribed hippocampal lesions using the same stimulus material and the same recognition procedures as were used in the two earlier studies (Westerberg et al., 2006; Holdstock et al., 2002). We then scored the data in two different ways, first as a 60-item yes/no test and a 12-item forced-choice test, as had been done in the two earlier studies; and second, as a more conventional 24-item yes/no test and a 12-item forced-choice test (by scoring only the first 24 test items of the yes/no test).

METHODS

Participants

Five memory-impaired patients participated (4 men) (Table 1), all of whom have bilateral lesions thought to be limited to the hippocampal region (CA fields, dentate gyrus, and subicular complex).

K.E. became amnesic in 2004 after an episode of ischemia associated with kidney failure and toxic shock syndrome. L.J. became amnesic in 1988 during a 6-month period with no known precipitating event. Her memory impairment has been stable since that time. Patients R.S. and G.W. became amnesic in 1998 and 2001, respectively, following a drug overdose and associated respiratory failure. Patient J.R.W. became amnesic in 1990 following an anoxic episode associated with cardiac arrest.

Estimates of medial temporal lobe damage were based on quantitative analysis of magnetic resonance images

Table 1. Characteristics of Amnesic Patients

				WMS—R				
Patient	*Age (years)*	*Education (years)*	*WAIS-III IQ*	*Attention*	*Verbal*	*Visual*	*General*	*Delay*
K.E.	64	13.5	108	114	64	84	72	55
L.J.	68	12	101	105	83	60	69	<50
R.S.	49	12	99	99	85	81	82	<50
G.W.	47	12	108	105	67	86	70	<50
J.R.W.	43	12	90	87	65	95	70	<50

The Wechsler Adult Intelligence Scale-III (WAIS-III) and the Wechsler Memory Scale—Revised (WMS—R) yield mean scores of 100 in the normal population, with a standard deviation of 15. The WMS—R does not provide numerical scores for individuals who score below 50. IQ scores for J.R.W. and R.S. are from the Wechsler Adult Intelligence Scale—Revised.

compared with data for 19 controls (for K.E., R.S., G.W., J.R.W), or 11 controls (L.J.). Nine coronal magnetic resonance images from these five patients are available as supplemental material to Wais et al. (2006). The volume of the full anterior–posterior length of the hippocampus and the parahippocampal gyrus were measured following published procedures (Insausti et al., 1998; Amaral & Insausti, 1990). For each patient, the hippocampal and parahippocampal gyrus volumes were divided by the intracranial volume to correct for brain size (Gold & Squire, 2005). Patients 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*s below the control mean). On the basis of two patients (L.M. and W.H.) with similar bilateral volume loss in the hippocampus for whom detailed postmortem neurohistological information was obtained (Rempel-Clower, Zola, Squire, & Amaral, 1996), this degree of volume loss likely reflects nearly complete loss of hippocampal neurons (Gold & Squire, 2005). In comparison, the volume of the parahippocampal gyrus is reduced by 17%, −8%, 1%, 12%, and 6%, respectively (all values within 2 *SD*s of the control mean).

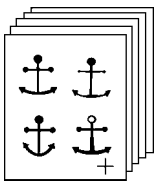
Additional measurements, based on four controls for each patient, were carried out for the frontal lobes, lateral temporal lobes, parietal lobes, occipital lobes, insular cortex, and fusiform gyrus (Bayley, Gold, Hopkins, & Squire, 2005). The only volume reduction in these regions greater than 1.3 *SD*s of the control mean was the parietal lobe of R.S. (Bayley et al., 2005). This finding for patient R.S. likely reflects natural variation in the parietal lobe volume rather than damage to this brain region as the parietal lobes are highly variable in size (Raz et al., 2005). In addition, no evidence of parietal lobe damage is apparent in the MRI scan of patient R.S. He also obtained a normal score on the block design test (scaled score = 11; WAIS-R, Wechsler, 1981), which is known to be sensitive to parietal lobe damage (Lezak, 1995).

The control group consisted of 15 individuals (11 men). Three controls were matched to each patient with respect to age (control mean = 58.0 years; range = 39–72; patient mean = 53.8 years), and years of education (control mean = 14.0 years; range = 12–18; patient mean = 12.3 years).

Materials

Four highly similar versions of each of 24 pictures of objects were used as stimuli, as provided to us courtesy of Andrew Mayes and Ken Paller (Table 2). Half of the pictures depicted living things, and half depicted man-made objects. As described by Holdstock et al. (2002), the stimuli were divided into two sets (set 1 and set 2) of 48 pictures each, which provided materials for two different test sessions. For the first test session, set 1 was used to construct a forced-choice test, and set 2 was used to construct a yes/no test. For the second test session, set 1 was used to construct the yes/no test and set 2 was used to construct the forced-choice test. These

Table 2. Organization of the Study and Test Phases for the Forced-choice Recognition Test (Top) and the Yes/No Recognition Test (Bottom)

			Test Phase						
						Number of Times a Target is Shown during Test			
Test Type	*Study Phase 12 Items*	*Delay*	*Images*	*Number of Cards*	*Images/Card*	*1×*	*2×*	*3×*	*Number of Foils*
Forced choice		45 sec		12	4	12	0	0	36
Yes/No				60	1	4	4	4	36

At study, 12 stimuli were presented, and the test was given 45 sec later. Altogether, there were two test sessions and each session tested both yes/no and forced-choice recognition memory using different stimuli. Each forced-choice test consisted of 12 trials in which a studied item was presented with three highly similar foils. Participants were asked to identify the target. In the example, the correct item is indicated by a "+". Each yes/no test consisted of 60 trials in which targets and foils were presented one at a time. There were 12 targets, 36 foils, and 12 additional trials in which four of the targets were repeated once and four other targets were repeated twice.

two versions of the test were given on separate days about 2 months apart (see Procedure).

Procedure

The procedure matched as closely as possible the procedure used previously with these materials (Holdstock et al., 2002). In each of two test sessions, we administered two recognition memory tests in immediate succession: yes/no and forced-choice. In one session, the forced-choice test was administered first followed by the yes/no test. In the other session, the yes/no test was given first, followed by the forced-choice test. Which session was scheduled first was counterbalanced across participants. The stimuli that were used for the forced-choice test in one session were used for the yes/no test in the other session. Likewise, the stimuli used for the yes/no test in one session were used for the forced-choice test in the other session. The two test sessions were separated by an average of 69 days (range = 51 to 124 days). Two of the controls were not available for a second test session, and their data were based on the results from the first session.

The study phase was identical for the yes/no and forced-choice tests (Table 2). Twelve stimuli were presented one at a time for 3 sec each, and participants made a living/man-made judgment for each picture in anticipation of a subsequent memory test. After all 12 pictures were presented, each picture was presented a second time for 3 sec each, and at this point it was emphasized that participants should try to memorize the details of each picture for the later test. Forty-five seconds after the study phase, participants were given either a yes/no or forced-choice recognition test (the retention interval was filled with simple arithmetic problems). During recognition testing, there was no time limit for responding. Upon completion of the first recognition test, the other recognition test was given (i.e., a new study phase and a new test).

For the yes/no recognition test, participants viewed one object at a time. Participants were asked to respond ''yes'' if the object was exactly the same as one viewed previously during the study phase. There were 60 test trials, consisting of the 12 targets, 36 foils (3 foils corresponding to each of the 12 targets), and 12 additional trials (4 of the targets were repeated once and 4 other targets were repeated twice) (Table 2). As described previously (Holdstock et al., 2002), the additional target items were included to minimize the possibility that participants would base their responses on how they had responded to similar items that had been presented earlier in the test. Thus, if a participant had already seen item A, he or she could not assume that any subsequent items that appeared to resemble item A would necessarily be foils. Following Holdstock et al. (2002), the score for the test was based on the response to the first occurrence of the 12 targets and on the response to the 36 foils. The score was virtually the same when we also calculated the score based on all occurrences of the target (see Results).

For the forced-choice test, there were 12 test trials. On each trial, participants viewed four items evenly spaced on a letter-sized page. Each display included the target item and three highly similar foils that corresponded to the target (Table 2). Participants were asked to identify the target item on each page.

Practice tests were given immediately before each study phase to familiarize participants with the study–test format. There were six novel study items and 30 novel test items. The pictures were of the same style as those used in the formal study. Following the procedure used by Holdstock et al. (2002), one practice test was given before the forced-choice test, and two practice tests were given before the yes/no recognition test.

Data Analysis

Results from the yes/no and forced-choice recognition tests were analyzed using a discriminability score (d') (Green & Swets, 1966). For the forced-choice data, d' was derived from Table A5.2 of MacMillan and Creelman (1991). The yes/no data were also examined more closely by calculating d' across data from only the first 24 trials, as well as across each of the five blocks of 12 trials. In these cases, a standard correction was sometimes necessary when calculating d' values, as the percent hits and the percent false alarms were sometimes 100% or 0%. Following MacMillan and Creelman, we converted 0% to $1/(2N)$% and 100% to $1 - 1/(2N)$% where N = the number of trials. For the analysis of the first 24 trials, this correction procedure was used once for one of the patients and once for one of the controls. For the analysis of each of the five blocks of 12 trials, the correction procedure was used, on average, 1.6 times per patient (range = 0–4) and, on average, 2.7 times per control (range = 0–4).

The scores obtained by the patients on the yes/no and forced-choice tests were also converted to z-scores. The z-score calculations for each test were based on the mean and standard deviation of the control scores.

RESULTS

Figure 1A shows the d' score of memory-impaired patients and controls on the yes/no and forced-choice recognition memory tests. Compared to controls, the patients exhibited impaired recognition memory performance on both tests. An analysis of variance (ANOVA) comparing the patients and controls revealed an effect of group [$F(1, 18) = 10.87, p < .01$] but no effect of test [$F(1, 18) = 1.13, p = .30$] and no Group × Test interaction [$F(1, 18) = 2.20, p = .16$]. Planned compar-

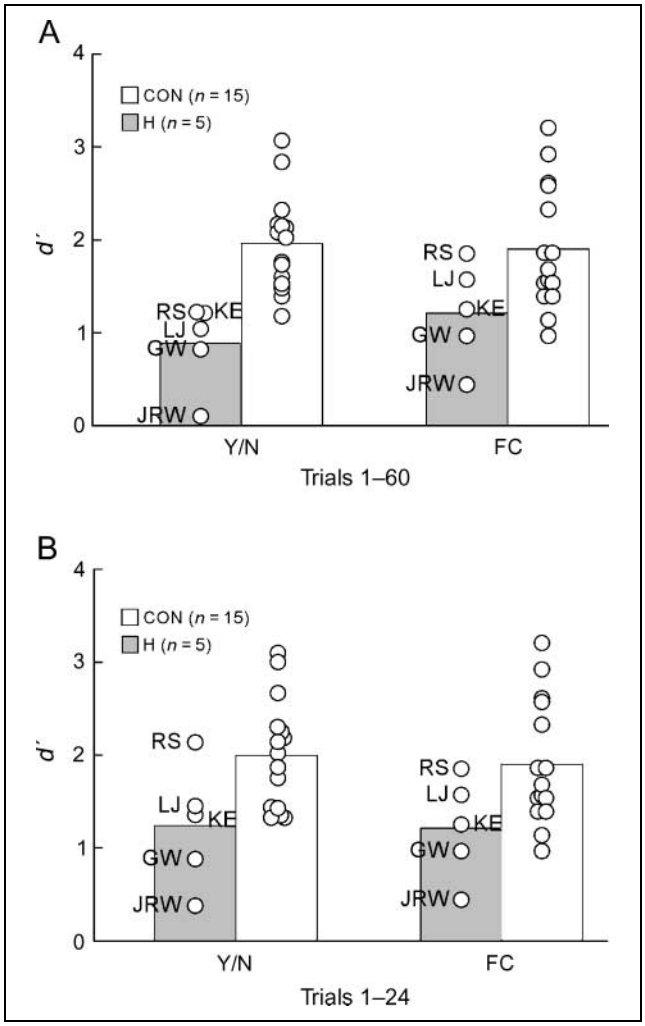


Figure 1. Discriminability performance (d') by patients with limited hippocampal lesions (H) and controls (CON) on yes/no recognition (Y/N) and forced-choice recognition (FC). For each test, participants first studied a set of 12 pictures (silhouettes of objects) and then took a recognition memory test. (A) The yes/no test consisted of 60 trials: the 12 targets (some of which appeared 2 or 3 times) and 36 foils. The forced-choice test consisted of 12 trials, each of which involved one studied picture and three very similar foils. The patients were impaired on both tests. Further, the yes/no score of the patients was lower than their forced-choice score ($p < .05$). (B) Performance on the first 24 trials of the yes/no object recognition test (Y/N) and on all 12 trials of the forced-choice object recognition test (FC). Again, the patients were impaired on both tests, but now the patients obtained similar scores on the yes/no and forced-choice tests. Each participant is represented by a circle, and patients are identified by their initials.

isons indicated that the patients were impaired on each test [yes/no: $t(18) = 4.14, p < .01$; forced-choice: $t(18) = 2.08, p < .05$]. Further, the patients, but not controls, performed more poorly on the yes/no test than on the forced-choice test [for patients, $d' = 0.88$ vs. 1.21; $t(4) = 2.93, p = .04$; for controls, $d' = 1.96$ vs. 1.91; $t(14) = .38, p = .71$]. For the patients on the yes/no test, mean hit rate = 71.7% and mean false alarm rate = 41.1%; for controls, mean hit rate = 85.4% and mean false alarm rate = 22.8%. The score was virtually the same when it was based on all occurrences of the target (for patients, $d' = 0.93$; for controls, $d' = 1.97$).

A potentially important difference between the yes/no and forced-choice tests is that there are 60 test trials in the yes/no test and only 12 trials in the forced-choice test. Consequently, the yes/no test takes longer to complete than the forced-choice test (the yes/no test took approximately 4 min to administer, and the forced-choice test took approximately 1.5 min), and the yes/no test may also generate more interference. In order to provide a stricter comparison between the two tests, we calculated performance on the yes/no test for only the first 24 test trials (i.e., the number of test trials that would appear in a conventional test of yes/no recognition memory following a presentation of 12 study items). On one version of the yes/no test, the first 24 trials consisted of 8 targets and 16 foils. On the second version of the yes/no test, the first 24 trials consisted of 8 targets (one of which was repeated) and 15 foils.

Figure 1B shows the performance of patients and controls on the first 24 trials of the yes/no recognition memory test. For comparison, the performance on the forced-choice test is also shown. An ANOVA comparing the d' scores of the patients and controls revealed an effect of group [$F(1, 18) = 6.55, p < .02$] but no effect of test [$F(1, 18) = .16, p = .69$] and no Group × Test interaction [$F(1, 18) = .06, p = .80$]. Planned comparisons indicated that the patients were impaired on each test [yes/no: $t(18) = 2.47, p < .03$; forced-choice: $t(18) = 2.08, p < .05$]. Notably, in contrast to the findings illustrated in Figure 1A, the patients performed similarly on the yes/no and forced-choice tests [for patients, $d' = 1.24$ vs. 1.21; $t(4) = .32, p = .77$; for controls, $d' = 2.00$ vs. 1.91; $t(14) = .57, p = .56$]. For the patients on the yes/no test, mean hit rate = 70.6% and mean false alarm rate = 31.0%; for controls, mean hit rate = 85.9% and mean false alarm rate = 20.8%. The score was virtually the same when it was based on all occurrences of the target (for patients, 1.17; for controls, 2.00). Thus, the patients performed as well on the yes/no test as on the forced-choice test when performance was assessed across the first 24 trials instead of across all 60 trials.

This difference in findings between Figure 1A and B was reflected in a marginally significant improvement in d' scores when performance on the yes/no test was assessed across 24 trials rather than all 60 trials [for 24 trials, $d' = 1.24$; for 60 trials, $d' = .87$; $t(4) = 2.36, p = .08$]. In contrast, the controls performed nearly the same across the first 24 trials as they did across all 60 trials [2.00 vs. 1.96; $t(14) = .52, p = .61$].

An analysis of z-scores also indicated that the patients performed more poorly on the yes/no test than on the forced-choice test when all 60 trials were scored [$z = -2.07$ vs. -1.03; $t(4) = 5.5, p < .01$] (as in Figure 1A). In contrast, the patients performed similarly on the two tests when only the first 24 trials were scored [$z = -1.30$ vs. -1.03; $t(4) = 1.63, p = .18$] (as in Figure 1B).

Further, a direct comparison of yes/no performance under the two scoring methods showed that performance of the patients was better across the first 24 trials than across all 60 trials [$z = -1.30$ vs. -2.07; $t(4) = 3.14$, $p = .04$].

In view of these findings, we next examined the performance of patients and controls across the entire yes/no recognition memory test by calculating the mean d' score on each block of 12 trials. As might be expected, the scores of the patients declined as testing continued. The d' scores were 1.07, 1.25, 0.80, 0.10, and 0.06 for blocks 1 to 5, respectively. The controls also declined somewhat (1.76, 2.10, 1.78, 1.09, 0.93) but the decline in patient performance was more marked.

An analysis of percent correct scores across blocks 1 to 5 revealed the same pattern. For patients, the scores were 67.8%, 71.5%, 59.4%, 50.0%, and 56.1% (chance = 50%). For controls, the scores were 80.3%, 82.2%, 80.8%, 70.0%, and 77.2%. All the control scores were above chance levels ($p < .001$). For the patients, only the scores for the first two blocks were above chance ($p < .01$).

DISCUSSION

Five patients with damage limited to the hippocampus took yes/no and forced-choice recognition memory tests that involved highly similar targets and foils. The patients were impaired on both kinds of test. Further, the patients were more impaired on the yes/no test than on the forced-choice test in apparent support of the suggestion that hippocampal damage might especially impair yes/no performance when the targets and foils are very similar (O'Reilly & Rudy, 2000). However, we also noted that the yes/no and forced-choice tests were not equivalent in a potentially important respect. Specifically, the two tests differed in the number of test trials that were given (60 trials in the yes/no test and 12 trials in the forced-choice test). As a result, the yes/no test took longer to complete than the forced-choice test and may also have generated more interference. In order to provide a more direct comparison between the two tests, we calculated performance on the first 24 test trials of the yes/no test (i.e., the number of test trials that would be given in a conventional test of yes/no recognition memory following the presentation of 12 study items). When this was done, the patients exhibited equivalent performance on the yes/no and forced-choice tests. Note that for the controls these two tests were equivalent in difficulty.

The results of the present study can be compared to two earlier studies in which the same recognition memory tests that we used were given to memory-impaired patients (Westerberg et al., 2006; Holdstock et al., 2002). Our results are in agreement with these two studies to the extent that when yes/no recognition performance was calculated across all 60 test trials, our patients were more impaired on the yes/no test than on the forced/choice test. One of these studies (Holdstock et al., 2002) involved a single patient (Y.R.), who is reported to have selective hippocampal damage. Compared to controls, Y.R. was impaired on the yes/no test ($d' = 0.94$) but scored within the range of the control scores on the forced-choice test ($d' = 1.55$). The second study (Westerberg et al., 2006) involved eight individuals with MCI whose memory dysfunction was attributed to preclinical Alzheimer's disease. Like Y.R., the MCI group was impaired on the yes/no test ($d' = 0.60$) but performed better on the forced-choice test ($d' = 0.93$).

It is interesting to note that the pattern of results reported in these two studies was evident in two of our own patients (patient L.J.—yes/no, $d' = 1.04$, forced-choice, $d' = 1.57$; patient R.S.—yes/no, $d' = 1.21$, forced-choice, $d' = 1.85$) (see Figure 1A). The magnitude of the difference between the scores on the two tests was similar to what was found for patient Y.R. (yes/no, $d' = 0.94$, forced-choice, $d' = 1.55$). Yet, although L.J. and R.S. performed close to the control mean on the forced-choice test, overall, our patient group was impaired on this test. These results illustrate the difficulty in generalizing from studies of only one or two patients.

It is also illuminating to consider the results obtained by the control groups in the two earlier studies. There is wide agreement that healthy individuals obtain similar scores on yes/no and forced-choice tests when performance is expressed as d' (Smith & Duncan, 2004; Kroll et al., 2002; Khoe et al., 2000; Yonelinas et al., 1992; Green & Moses, 1966). The control group in the present study also exhibited this pattern of results. In contrast, the control groups in the two earlier studies (Westerberg et al., 2006; Holdstock et al., 2002) exhibited noticeably poorer performance on the forced-choice test than on the yes/no test. These unexpectedly low scores on the forced-choice test contributed to the finding that the patients in those studies did not appear impaired on the forced-choice test.

The main finding of our study was that the patients performed equivalently on the forced-choice and yes/no tests when just the first 24 trials of the yes/no test were scored. The question naturally arises as to whether this same finding might have been obtained in the two earlier studies (Westerberg et al., 2006; Holdstock et al., 2002). In fact, Westerberg et al. (2006) did report performance across the first and second half of their yes/no test (first 30 trials and second 30 trials). Consistent with what we observed in our patients, the scores of the MCI patients declined across the yes/no test session (first half of the yes/no test vs. second half; $d' = 0.97$ vs. 0.57). In contrast, the scores of the control group numerically increased (1.29 vs. 1.39). As a result, the scores of the MCI patients on the first half of the yes/no test closely matched their forced-choice test scores ($d' = 0.97$ vs. 0.93) and were also not significantly impaired rela-

tive to controls scores. The difference between yes/no performance and forced-choice performance emerged only in the second half of the 60-trial test ($d' = 0.57$ vs. 0.93). Certainly, the length of the yes/no test needs to be taken into consideration when comparing yes/no and forced-choice recognition performance. Possible explanations for why yes/no performance declined when the test was long include the sensitivity of patients to a long study–test delay, the build-up of interference caused by multiple presentations of targets, and the build-up of interference caused by multiple presentations of foils. In a footnote, Westerberg et al. reported that a new group of MCI patients obtained impaired yes/no recognition scores even in the first half of the recognition test (similar to our findings for the first 24 trials of the 60-trial yes/no test). Unlike what was observed in their main MCI group, the scores of the new MCI group did not decline across the yes/no test session. It was not reported how yes/no performance compared to forced-choice performance.

These conclusions about the disadvantages of a 60-trial yes/no test are based on our analysis of the data from the first 24 trials of the test. A possible limitation of this 24-trial analysis is that there were not an equal number of targets and foils (8 or 9 targets and 15 or 16 foils in the two different versions of the test). Future studies could determine more precisely how performance changes across lengthy yes/no tests in memory-impaired patients and controls and how yes/no performance compares to forced-choice performance. It is also possible that a difference between yes/no performance might emerge if the tests were given under conditions different from the conditions of our study (e.g., long study-test delays).

The purpose of our study was to test the suggestion that patients with hippocampal damage might be impaired on yes/no recognition memory tests but be intact on forced-choice tests when targets and foils are very similar (O'Reilly & Norman, 2002; O'Reilly & Rudy, 2000). In the yes/no test, only one stimulus is presented at a time. Under these circumstances, it was suggested that discriminating between targets and foils is very difficult and that participants can benefit by engaging in recollection. Because patients with hippocampal damage are proposed to be impaired at recollection, they would be expected to perform poorly on the yes/no test. In contrast, in the forced-choice test, targets and foils are presented simultaneously. Under these circumstances, it was suggested that studied items will yield a small but reliably stronger familiarity signal than the foils. Because patients with hippocampal damage are proposed to have intact familiarity, they would be expected to perform well on the forced-choice test.

The results of the present study did not support these expectations. First, forced-choice performance was impaired in our patients. Second, when a strict comparison was made between forced-choice and yes/no scores (by scoring the first 24 trials of the yes/no test), the patients performed similarly on the two tests. There appear to be two ways to understand this result. One possibility is that forced-choice and yes/no tests do indeed depend differentially on recollection and familiarity but that the hippocampus supports both of these processes (Rutishauser et al., 2006; Wais et al., 2006). Accordingly, hippocampal damage impaired performance on the two tests similarly. A second possibility is that forced-choice and yes/no tests depend similarly on the recollection and familiarity components of recognition memory. Accordingly, patients with hippocampal damage performed similarly on these two tests. Indeed, by this scenario, their performance would not be expected to differ on the two tests, regardless of the extent to which recollection and familiarity are differentially impaired.

One way to pursue these questions would be to test patients who have poorer recall than recognition, such as patients with frontal lesions (Wheeler, Stuss, & Tulving, 1995). Recall is typically thought to depend on recollection, whereas recognition depends on both recollection and familiarity. Accordingly, if some yes/no tests depend on recollection more than corresponding forced-choice tests do, then patients with frontal lesions should be disadvantaged at yes/no recognition.

In summary, the present study found that memory-impaired patients with damage to the hippocampus were impaired on tests of recognition memory in which the targets and foils were very similar. The patients initially appeared to perform more poorly on the yes/no test than on the forced-choice test. However, further inspection of the data revealed that this disadvantage for the patients on the yes/no test was due to the fact that their performance declined across the 60 trials of the test. When this decline was accounted for, and the yes/no test was scored in the conventional fashion, the patients performed virtually identically on the yes/no and forced-choice tests. The findings do not encourage the view that yes/no and forced-choice recognition tests can be used to illuminate the functions of the hippocampus.

Acknowledgments

This study was supported by the Medical Research Service of the Department of Veterans Affairs, The National Institute of Mental Health (MH24600), and the Metropolitan Life Foundation. We thank Jennifer Frascino for research assistance and Andrew Mayes and Ken Paller for providing test materials.

Reprint requests should be sent to L. R. Squire, Department of Psychiatry, University of California, San Diego, San Diego, CA 92093, or via e-mail: lsquire@ucsd.edu.

REFERENCES

Aggleton, J., & Shaw, C. (1996). Amnesia and recognition memory: A re-analysis of psychometric data. *Neuropsychologia, 34,* 51–62.

Amaral, D. G., & Insausti, R. (1990). Hippocampal formation. In G. Paxinos (Ed.), *The human nervous system*. San Diego: Academic Press.

Bastin, C., & Van der Linden, M. (2003). The contribution of recollection and familiarity to recognition memory: A study of the effects of test format and aging. *Neuropsychology, 17,* 14–24.

Bayley, P. J., Gold, J. J., Hopkins, R. O., & Squire, L. R. (2005). The neuroanatomy of remote memory. *Neuron, 46,* 799–810.

Brown, M. W., & Aggleton, J. P. (2001). Recognition memory: What are the roles of the perirhinal cortex and hippocampus? *Nature Reviews Neuroscience, 2,* 51–61.

Gold, J. J., & Squire, L. R. (2005). Quantifying medial temporal lobe damage in memory-impaired patients. *Hippocampus, 15,* 79–85.

Green, D. M., & Moses, F. L. (1966). On the equivalence of two recognition measures of short-term memory. *Psychological Bulletin, 66,* 228–234.

Green, D. M., & Swets, J. A. (1966). *Signal detection theory and psychophysics*. New York: Wiley.

Holdstock, J. S., Mayes, A. R., Roberts, N., Cezayirli, E., Isaac, C. L., O'Reilly, R. C., et al. (2002). Under what condition is recognition spared relative to recall after selective hippocampal damage in humans? *Hippocampus, 12,* 341–351.

Insausti, R., Juottonen, K., Soininen, H., Insausti, A. M., Partanen, K., Vainio, P., et al. (1998). MR volumetric analysis of the human entorhinal, perirhinal, and temporopolar cortices. *American Journal of Neuroradiology, 19,* 659–671.

Khoe, W., Kroll, N. E., Yonelinas, A. P., Dobbins, I. G., & Knight, R. T. (2000). The contribution of recollection and familiarity to yes–no and forced-choice recognition tests in healthy subjects and amnesics. *Neuropsychologia, 38,* 1333–1341.

Kroll, N. E., Yonelinas, A. P., Dobbins, I. G., & Frederick, C. M. (2002). Separating sensitivity from response bias: Implications of comparisons of yes–no and forced-choice tests for models and measures of recognition memory. *Journal of Experimental Psychology. General, 131,* 241–254.

Lezak, M. D. (1995). *Neuropsychological assessment* (3rd ed.). New York: Oxford University Press.

MacMillan, N. A., & Creelman, C. D. (1991). *Detection theory: A users guide*. Cambridge: Cambridge University Press.

Mandler, G. (1980). Recognizing: The judgment of previous occurrence. *Psychological Review, 87,* 252–271.

Mayes, A. R., Holdstock, J. S., Isaac, C. L., Hunkin, N. M., & Roberts, N. (2002). Relative sparing of item recognition memory in a patient with adult-onset damage limited to the hippocampus. *Hippocampus, 12,* 325–340.

O'Reilly, R. C., & Norman, K. A. (2002). Hippocampal and neocortical contributions to memory: Advances in the complementary learning systems framework. *Trends in Cognitive Sciences, 6,* 505–510.

O'Reilly, R. C., & Rudy, J. W. (2000). Computational principles of learning in the neocortex and hippocampus. *Hippocampus, 10,* 389–397.

Parkin, A. J., Yeomans, J., & Bindschaedler, C. (1994). Further characterization of the executive memory impairment following frontal lobe lesions. *Brain and Cognition, 26,* 23–42.

Raz, N., Lindenberger, U., Rodrigue, K. M., Kennedy, K. M., Head, D., Williamson, A., et al. (2005). Regional brain changes in aging healthy adults: General trends, individual differences and modifiers. *Cerebral Cortex, 15,* 1676–1689.

Rempel-Clower, N. L., Zola, S. M., Squire, L. R., & Amaral, D. G. (1996). Three cases of enduring memory impairment after bilateral damage limited to the hippocampal formation. *Journal of Neuroscience, 16,* 5233–5255.

Rutishauser, U., Mamelak, A. N., & Schuman, E. M. (2006). Single-trial learning of novel stimuli by individual neurons of the human hippocampus–amygdala complex. *Neuron, 49,* 805–813.

Smith, D. G., & Duncan, M. J. (2004). Testing theories of recognition memory by predicting performance across paradigms. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 30,* 615–625.

Squire, L. R., Wixted, J. T., & Clark, R. E. (2007). Recognition memory and the medial temporal lobe: A new perspective. *Nature Reviews Neuroscience, 8,* 872–883.

Wais, P. E., Wixted, J. T., Hopkins, R. O., & Squire, L. R. (2006). The hippocampus supports both the recollection and the familiarity components of recognition memory. *Neuron, 49,* 459–466.

Wechsler, D. (1981). *Wechsler Adult Intelligence Scale—Revised*. New York: Psychological Corporation.

Westerberg, C. E., Paller, K. A., Weintraub, S., Mesulam, M. M., Holdstock, J. S., Mayes, A. R., et al. (2006). When memory does not fail: Familiarity-based recognition in mild cognitive impairment and Alzheimer's disease. *Neuropsychology, 20,* 193–205.

Wheeler, M. A., Stuss, D. T., & Tulving, E. (1995). Frontal lobe damage produces episodic memory impairment. *Journal of the International Neuropsychological Society, 1,* 525–536.

Wixted, J. T. (2007). Dual-process theory and signal-detection theory of recognition memory. *Psychological Review, 114,* 152–176.

Wixted, J. T., & Squire, L. R. (2004). Recall and recognition are equally impaired in patients with selective hippocampal damage. *Cognitive, Affective & Behavioral Neuroscience, 4,* 58–66.

Yonelinas, A. P., Hockley, W. E., & Murdock, B. B. (1992). Test of the list-strength effect in recognition memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 18,* 345–355.

Yonelinas, A. P., Kroll, N. E., Quamme, J. R., Lazzara, M. M., Sauve, M. J., Widaman, K. F., et al. (2002). Effects of extensive temporal lobe damage or mild hypoxia on recollection and familiarity. *Nature Neuroscience, 5,* 1236–1241.