Behavioral Neuroscience
1996, Vol. 110, No. 5, 861–871

Dissociable Properties of Memory Systems: Differences in the Flexibility of Declarative and Nondeclarative Knowledge

Paul J. Reber and Barbara J. Knowlton
University of California, San Diego

Larry R. Squire
Veterans Affairs Medical Center
and University of California, San Diego

Amnesic patients (n = 8), who have severely impaired declarative memory, learned a probabilistic classification task at the same rate as normal subjects (n = 16) but subsequently were impaired on transfer tests that required flexible use of their task knowledge. A second group of controls (n = 20) rated the questions on the transfer tests according to whether the questions simply reinstated the training conditions or required flexible use of task knowledge. The amnesic patients tended to be impaired on the same items that were rated as requiring indirect or flexible use of knowledge. Thus, control subjects acquired declarative knowledge about the task that could be applied flexibly to the transfer tests. The nondeclarative memory available to amnesic patients was relatively inflexible and available only in conditions that reinstantiated the conditions of training. These findings show that declarative memory has different operating characteristics than nondeclarative memory.

Neurobehavioral studies of humans and experimental animals with selective lesions have distinguished between a kind of memory dependent on medial temporal lobe and diencephalic brain structures (declarative memory) and various other kinds of (nondeclarative) memory that are independent of these structures (Mishkin & Petri, 1984; Schacter & Tulving, 1994a; Squire & Zola-Morgan, 1991). It has been noted that separate memory systems should not be founded solely on lesion evidence or any particular dissociation, because the dissociations are so closely tied to particular experimental procedures (Roediger, 1990). Presumably, memory systems that are distinct from each other should be separable on the basis of multiple criteria (Nadel, 1994; Schacter, 1992; Sherry & Schacter, 1987; Weiskrantz, 1991). Thus, it might be possible to distinguish declarative and nondeclarative memory systems not only in terms of anatomy, but also in terms of operating characteristics, the kind of information processed, and the purpose served by each system (Tulving, 1991). What is needed is a property list for declarative and nondeclarative memory systems, independent of evidence concerning what humans or experimental animals with specific lesions can and cannot learn (Schacter & Tulving, 1994b; Squire, 1994).

One property that has been explored in recent studies is the awareness that subjects have for what has been learned. Whereas declarative memory affords the capacity for conscious recollections about facts and events, nondeclarative memory does not appear to require awareness of any memory content. Several tasks (e.g., artificial grammar learning, sequence learning, and some kinds of category learning) are performed normally by amnesic patients despite their inability to remember the events of training (Knowlton, Ramus, & Squire, 1992; Knowlton, Squire, & Gluck, 1994; Nissen & Bullemer, 1987; P.J. Reber & Squire, 1994). Information appears to be acquired implicitly without requiring awareness of what has been learned (A.S. Reber, 1989; P.J. Reber & Squire, 1994; Willingham, Nissen, & Bullemer, 1989; for criticism and discussion of this conclusion, see Shanks & St. John, 1994, and associated commentary). Although awareness is an important characteristic of declarative memory and central to its formulation, it is a subjective notion and one that has little utility when trying to make links between memory systems in humans and nonhuman animals.

A second property that may be useful in distinguishing declarative and nondeclarative memory concerns the flexibility of stored knowledge. Declarative memory is considered to support the flexible use of stored knowledge so that task knowledge can be used in situations different from the original learning context. Nondeclarative memory, in contrast, has been considered to be more closely tied to the original learning situation and less accessible to other systems (Cohen, 1984; Squire, 1994). Three experiments with rats and monkeys have identified apparent differences in the flexibility of declarative and nondeclarative memory. In the first study (Eichenbaum, Mathews, & Cohen, 1989), rats with fornix lesions learned to select the rewarded odor in each of two repeatedly presented odor pairs (e.g., A+B−, C+D−). Having learned the discriminations, the fornix-lesioned rats were markedly impaired compared with controls when the rewarded odors were paired with different incorrect odors (e.g., A+D−, C+B−). In the second study (Eichenbaum, Stewart, & Morris, 1990), rats with

Paul J. Reber and Barbara J. Knowlton, Department of Psychiatry, University of California, San Diego; Larry R. Squire, Veterans Affairs Medical Center and Departments of Psychiatry and Neurosciences, University of California, San Diego. Barbara J. Knowlton is currently at the Department of Psychology, University of California, Los Angeles.

We thank N. Champagne, B. Kronenberg, and J. Zouzounis for research assistance. This work was supported by the Medical Research Service of the Department of Veterans Affairs, National Institute of Mental Health (NIMH) Grant MH24600, an NIMH Postdoctoral fellowship, the McKnight Foundation, and the McDonnell-Pew Center for Cognitive Neuroscience, San Diego.

Correspondence concerning this article should be addressed to Larry R. Squire, Veterans Affairs Medical Center (116A), 3350 La Jolla Village Drive, San Diego, California 92161.

fornix lesions learned to swim to an invisible platform in a circular pool of opaque water when the same starting point was used on each trial. However, the same animals were markedly impaired when they were required to swim from a novel starting point. Finally, Saunders and Weiskrantz (1989) trained monkeys with lesions of the hippocampus or related structures to discriminate between rewarded object pairs (AB+ and CD+) and nonrewarded object pairs (AC− and BD−). The monkeys with lesions were subsequently impaired in comparison with control monkeys in tasks that asked them to identify the objects that composed the rewarded pairs (e.g., present A, then provide a choice between Object B and Object C). In each study, the impairment in the lesioned animals appeared to reflect an inability to use task knowledge flexibly. The lesioned animals acquired the ability to do the tasks but could not perform transfer tests when the testing situation varied from the original learning situation. Control animals presumably acquired more declarative knowledge about the tasks than the lesioned animals. This information may have supported flexible performance, whereas lesioned animals had to rely largely on less flexible, nondeclarative memory.

In studies of humans, it has been difficult to find clear evidence of inflexibility in the operation of nondeclarative memory. Glisky, Schacter, and Tulving (1986a, 1986b) found that amnesic patients were able to learn a rather complex task after very extensive training (commands for operating a computer). The knowledge they acquired was "hyperspecific" in the sense that the patients were unable to use their knowledge in a novel situation or to answer questions when the wording was different than it was during training. However, in view of the extensive training required by the patients, it is unclear whether impaired transfer performance was due to inflexibility of nondeclarative memory or inflexibility of the residual declarative memory that was acquired abnormally slowly during extended training. Indeed, this same uncertainty applies to the studies with rats just reviewed, because the lesioned animals required significantly more training to learn the tasks initially (Eichenbaum et al., 1989, 1990).

A different study suggested that residual declarative memory acquired by amnesic patients is as flexible as the declarative memory ordinarily acquired by normal subjects (Shimamura & Squire, 1988). In this study, amnesic patients and control subjects learned sentences during four study trials, each of which was followed by a cued recall test. The patients learned much less than the control subjects, consistent with the idea that cued recall depends substantially on declarative memory. However, when cued-recall performance was equated between amnesic patients and controls (by delaying the testing of controls for 1 to 2 weeks), the amnesic patients were as confident of their answers as control subjects and performed as well as control subjects on a transfer test that used indirect cues. This finding showed that the residual (presumably declarative) knowledge retained by amnesic patients is as flexible, as accessible to indirect cues, and as available to awareness as the knowledge retained by control subjects. The same conclusion was reached in a second study (Butters, Glisky, & Schacter, 1993), which found that as the cued-recall performance of amnesic patients improved with extended training, so did their ability to respond to indirect cues. These studies suggest that inflexibility, if and when it is observed, is likely to be a property of nondeclarative memory, not a property of gradually acquired declarative memory.

To study directly whether nondeclarative memory is inflexible, one must assess performance on transfer tests in a case where the original learning is supported substantially by nondeclarative memory and declarative memory is essentially unavailable. One way to establish this condition is to identify a task that amnesic patients learn at the same rate as control subjects. Amnesic patients have severely impaired declarative memory but perform normally on a variety of tasks that depend on nondeclarative (implicit) memory. The performance of amnesic patients on transfer tests should then reflect the properties of nondeclarative memory.

Recently, amnesic patients and control subjects were found to acquire a probabilistic classification task at the same rate, at least during the first 50 trials (Knowlton et al., 1994). For the present study, we designed three transfer tests, involving a total of 16 questions, to assess the ability of subjects to use their task knowledge flexibly. To perform the transfer tasks, the amnesic patients would need to depend primarily on their nondeclarative knowledge of the task and on what little declarative knowledge they could acquire. By contrast, control subjects would have been able in parallel to acquire declarative knowledge about the task while they acquired the ability to perform it, and this knowledge could potentially be used in transfer tests. Accordingly, poorer performance by the amnesic patients on the transfer tasks, in comparison with control subjects, would imply that the nondeclarative knowledge that amnesic patients rely on to learn the task is relatively inflexible.

Experiment 1

Method

Participants

Amnesic patients. Eight amnesic patients (6 men and 2 women) participated in this study. Three of the patients have Korsakoff's syndrome. All 3 have participated in quantitative magnetic resonance imaging (MRI) studies that demonstrated reductions in the volume of the mammillary nuclei (for patients R.C., J.W., and P.N., see Squire, Amaral, & Press, 1990). Patient M.G. sustained a bilateral medial thalamic infarction, which was confirmed by MRI (unpublished observations). Of the remaining 4 patients, 3 have bilateral reduction in the size of the hippocampal formation confirmed by MRI (for P.H., see Polich & Squire, 1993; for L.J. and H.W., unpublished observations). The final patient (A.B.) is unable to participate in MRI studies, but the etiology of his amnesia (anoxia) is consistent with hippocampal damage. All 8 patients are well characterized neuropsychologically (see Tables 1 and 2).

The patients averaged 65.9 years of age at the beginning of the study and had an average of 13.8 years of education. Immediate and delayed (12 min) recall of a short prose passage averaged 6.8 and 0 segments, respectively (Gilbert, Levee, & Catalano, 1968; maximum number of segments = 21). Scores on other memory tests appear in Tables 1 and 2. The mean score on the Dementia Rating Scale was 132.1 (Mattis, 1976; maximum score = 144). Most of the points that were lost were on the memory subportion of the test (mean points lost = 6.8). The mean score for the Boston Naming Test was 56 (Kaplan, Goodglass, & Weintraub, 1983; maximum score = 60). Scores for normal subjects on these tests can be found elsewhere (Janowsky, Shimamura, Kritchevsky, & Squire, 1989; Squire et al., 1990).

Table 1
Characteristics of Amnesic Patients

Patient	Lesion	Age (years)	WAIS–R IQ	WMS–R Attention	WMS–R Verbal	WMS–R Visual	WMS–R General	WMS–R Delay
A.B.	HF	57	104	87	62	72	54	<50
P.H.	HF	72	115	117	67	83	70	57
L.J.	HF	57	98	105	83	60	69	<50
H.W.	HF	76	109	97	84	102	89	54
R.C.	Dien	78	106	115	76	97	80	72
P.N.	Dien	67	99	81	77	73	67	53
J.W.	Dien	58	98	104	65	70	57	57
M.G.	Dien	62	97	92	97	77	89	72
M		65.9	103.2	99.8	76.4	79.2	71.9	58.1

Note. WAIS–R = Wechsler Adult Intelligence Scale—Revised; WMS–R, Wechsler Memory Scale—Revised; HF = hippocampal formation; Dien = diencephalon. The WAIS–R and the WMS–R indexes yield a mean score of 100 in the normal population with a standard deviation of 15. The WMS–R does not provide scores for subjects who score below 50. Therefore, the two scores below 50 were scored as 50 for calculating a group mean.

Control subjects. The control subjects were either employees or volunteers at the San Diego Veterans Affairs (VA) Medical Center or were recruited from the retirement community of the University of California, San Diego (UCSD). The group consisted of 16 subjects (7 men and 9 women) matched to the amnesic patients with respect to the mean and range of their ages, years of education, and scores on the Information and Vocabulary subtests of the Wechsler Adult Intelligence Scale—Revised (WAIS–R). They averaged 66.1 years of age (range = 53–77), 15.2 years of education, and 21.7 and 52.0 on the Information and Vocabulary subtests, respectively (amnesic patients = 19.5 and 54.5). Immediate and delayed recall of the short prose passage averaged 7.4 and 6.1 segments, respectively.

Materials and Procedure

The weather prediction task. Subjects were given a probabilistic classification task using a Macintosh Powerbook computer (Knowlton et al., 1994). The task required subjects to decide on each trial which of two weather outcomes (rainy or sunny weather) would occur on the basis of a set of one, two, or three cues (out of four possible cues; see Figure 1). Each of the four cues was independently associated to each outcome with a fixed probability, and the two outcomes occurred equally often. There were four possible cue–outcome association strengths: A cue was associated either 77%, 58%, 42%, or 23% (approximately) with a particular weather outcome.

Table 3 shows the distribution of outcomes for the 14 possible combinations of cues that were presented during a 50-trial learning test (the case of all four cues present and the case of no cues present did not occur). Within both the control subject and amnesic patient groups, each cue was equally likely to be assigned one of the four association strengths. There were 4! (4 factorial) or 24 possible ways in which the cues could be assigned their association strengths.

On each trial, one, two, or three of the four possible cues were presented. Subjects responded by predicting a weather outcome (rain or sunshine) using the computer keyboard on which two adjacent response keys were marked with icons representing sun and rain. After each prediction was made, the computer screen displayed the actual weather outcome. In addition, if the prediction was correct, a

Table 2
Memory Test Performance

Patient	Diagram recall	Paired associates			% word recall	% word recognition	50 words	50 faces
A.B.	4	1	1	2	33	83	32	33
P.H.	3	0	0	1	27	84	36	34
L.J.	3	0	0	0	40	93	33	29
H.W.	6	0	1	0	31	85	23	22
R.C.	3	0	0	0	19	85	37	30
P.N.	2	1	1	1	29	83	31	31
J.W.	4	0	0	2	28	96	29	34
M.G.	6	1	2	8	52	89	30	34
M	3.9	0.38	0.62	2.12	32.4	87.2	31.4	30.9
Control *M* (n = 8)	20.6	6.0	7.6	8.9	71.0	97.0	41.1	38.1

Note. The diagram recall score is based on delayed (12-min) reproduction of the Rey–Osterrieth figure (Osterrieth, 1944; maximum score = 36). The average score for the amnesic patients for copying the figure was 28.0, a normal score (Kritchevsky, Squire, & Zouzounis, 1988). The paired associate scores are the number of word pairs recalled on three successive trials (maximum score = 10/trial). The word recall score is the percentage of words recalled across five successive study–test trials. The word recognition score is the percentage of words identified correctly by yes–no recognition across five successive study–test trials. The score for words and faces is based on a 24-hr recognition test of 50 words or 50 faces (modified from Warrington, 1984; maximum score = 50, chance = 25). The mean scores for control subjects shown for these tests are from Squire and Shimamura (1986).

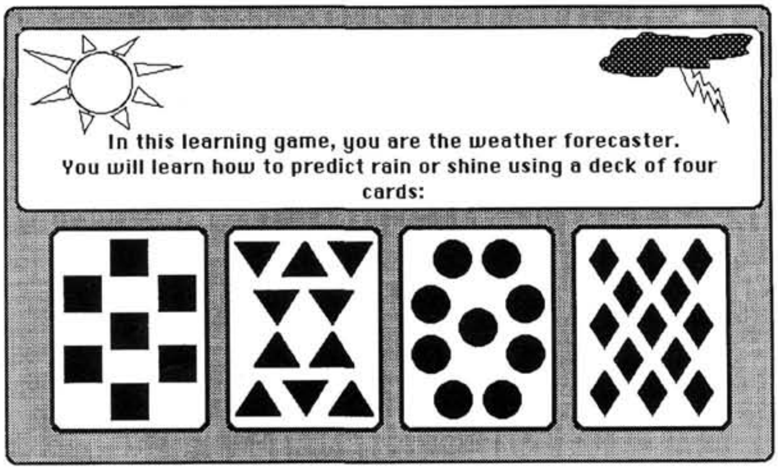

Figure 1. The task as it appeared on the computer screen. The icons indicating each weather outcome are displayed above the four cues.

high-pitched tone sounded, and a smiling face appeared on the screen. If the prediction was incorrect, a low-pitched tone sounded, and a frowning face appeared on the screen. If no prediction was made within 5 s, the weather outcome was shown and feedback was given as if an incorrect prediction had been made. (These missed trials were a 0.5-not scored.) A vertical bar at the right of the screen, set initially at 600, increased by one unit for each correct response and decreased by one unit for each incorrect response. The feedback remained on the screen for 2 s and was followed by a 0.5-s intertrial interval.

Seven of the amnesic patients (L.J., A.B., H.W., M.G., P.H., R.C., and J.W.) had been given the weather prediction task one time ($n = 5$) or two times ($n = 2$) on previous occasions at least 4 months earlier

Table 3
Probability Structure of the Weather Prediction Task

Cue				Frequency	Sunshine	Rain	*P*(Sun)
1	1	0	0	4	4	0	1.00
1	1	1	0	1	1	0	1.00
1	0	0	0	7	6	1	.86
1	0	1	0	4	3	1	.75
0	1	0	0	5	3	2	.60
1	0	0	1	2	1	1	.50
0	1	1	0	2	1	1	.50
1	0	1	1	2	1	1	.50
1	1	0	1	2	1	1	.50
0	0	1	0	5	2	3	.40
0	1	0	1	4	1	3	.25
0	0	0	1	7	1	6	.14
0	1	1	1	1	0	1	.00
0	0	1	1	4	0	4	.00
Total				50	25	25	

Note. Each row refers to 1 of the 14 possible trial types used during 50 training trials. The numbers under the Cue column refer, from left to right, to the cue most associated with sun; the cue second most associated with sun; the cue second most associated with rain; and the cue most associated with rain: 1 indicates that the cue was present on the trial; 0 indicates that the cue was absent on the trial. Thus, 1100 denotes a trial on which two cues were presented, the one most associated with sun and the one next most associated with sun. Frequency indicates the frequency with which each cue combination occurred. Sunshine and Rain indicate how often the cue combination resulted in each outcome. The final column indicates the probability that the weather outcome was sunny for each cue combination. *P*(Sun) = the probability of sun; *P*(rain) = 1 − *P*(sun).

(mean delay = 13.4 months, range = 4–22). Because each test session involved a different assignment of cue cards to weather outcomes, any residual knowledge from earlier test sessions would be expected to interfere with the present learning test rather than to facilitate it.

Transfer tests: Estimates of association strength. Immediately after subjects had completed 50 learning trials, we administered three transfer tests to assess their ability to flexibly use the knowledge they had acquired about the relationship between the four cues and the weather outcome. For each test, the cues were presented to subjects on separate printed cards that were approximately the same size as the cues on the computer screen. A printed card of instructions was in view throughout the first two tests.

1. *Single cues:* Subjects were shown each of the four cues one at a time and were asked in turn, "If just this card is showing, what percentage of the time will it be rainy [sunny]?" The weather outcome used in the question was consistent for each subject and balanced across subjects (i.e., half the subjects were asked all four questions with respect to rainy weather, and the other half were asked all four questions with respect to sunny weather). Subjects were instructed to respond with a number from 0 to 100.

2. *Cue pairs:* Subjects were shown each of the six possible pairs of cues one at a time and were asked in turn, "If these two cards are showing, what is the percentage of the time that it will be rainy [sunny]?". As in Number 1 above, the specific weather outcome that appeared in the six questions was consistent within subjects and balanced across subjects. Subjects responded with a number from 0 to 100.

3. *Cue selection:* Subjects were shown all four cues and were asked, "What if you knew it was going to be rainy and one card was showing? Which card would be most likely to be showing?" The same question was then repeated for the sunny weather outcome. Next, the same two questions were asked except that subjects were asked which *two* cues were most likely to be showing. Finally, the same two questions were asked, and subjects were asked which *three* cues were most likely to be showing. In response to each of these six questions, subjects identified the appropriate number of cards from the display. The left-to-right arrangement of the cards was randomized (shuffled manually by the experimenter) after each question. Half the subjects were asked about the rainy outcome first and then about the sunny outcome, and half were asked the questions in the reverse order.

Questionnaire. After completing the three tests of task knowledge, subjects were given eight four-alternative, multiple-choice questions to determine how well they remembered the testing situation. These questions asked, for example, about the layout of the screen, the number of cards that could appear together on the computer screen, the number of weather prediction trials presented, and the appearance of the cues.

Data analysis. For scoring weather prediction performance, a subject was considered to have made a correct response on a particular trial if the subject selected the outcome that was most associated with the cue pattern. Thus, subjects could have been scored as making a correct response (because they selected the most likely outcome), even though on that particular trial the feedback they received told them that their response was incorrect. In this way, the percent correct score reflected how well subjects learned the probabilistic associations between the cues and the two outcomes. Percent correct scores were analyzed in blocks of 10 trials for the 50 trials of the learning test. Within each block of 10 trials, percent correct scores were calculated separately for the two weather outcomes (rain and sunshine), and these scores were averaged together for the block score. Chance performance was 50%. Trials in which the two outcomes were equally likely (on 8 trials out of 50, distributed evenly across the test session; see Table 3, middle rows) were not included in the data analysis, because a subject's choice on these trials provided no information about classification learning.

Results

Weather Prediction

Figure 2 shows performance on the weather prediction task during 50 learning trials. Amnesic patients missed an average of 1.1 trials, and control subjects missed an average of 1.4 trials (i.e., no response was made within 5 s). Performance of the amnesic patients and control subjects was compared using a 2 (groups) × 5 (trial blocks) analysis of variance (ANOVA). There was no effect of group, $F(1, 22) < 1.0$; a marginal effect of block, $F(1, 22) = 3.68$, $p < .07$; and no interaction, $F(1, 22) < 1.0$. Although the improvement in performance across blocks did not reach significance ($p < .07$), both groups exhibited task knowledge by performing significantly above chance on the final 10-trial block, $t(7) = 3.29$, $p < .02$ for the amnesic patients and $t(15) = 3.76$, $p < .01$ for the control subjects. The two groups did not differ significantly from each other on any block ($ts < 1.0$), and they performed near chance level (50% correct) on the first block of trials ($ts < 1.0$).

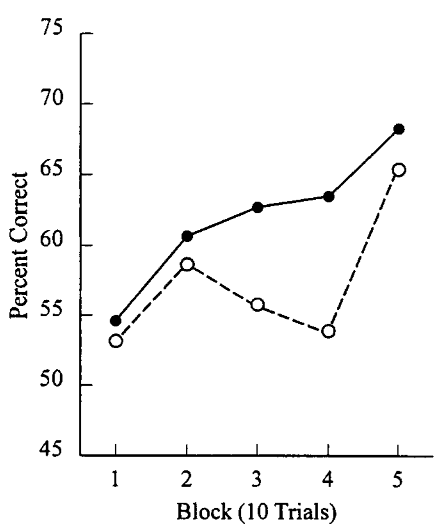

Figure 2. Percent correct performance for control subjects (filled circles) and amnesic patients (open circles) during 50 learning trials (10 trials/block). Chance performance is 50% correct. The standard errors of the mean ranged from 4.6% to 8.4%.

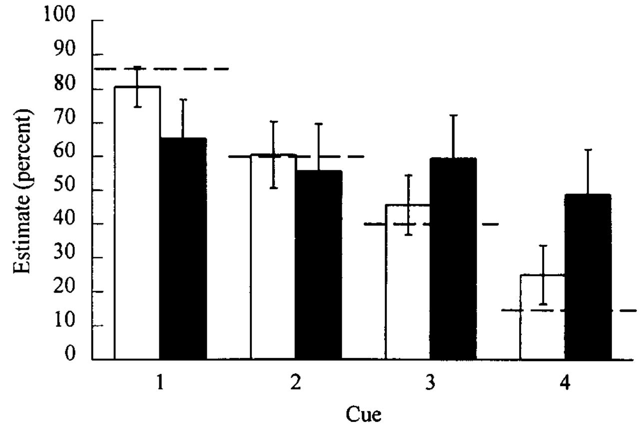

Figure 3. Estimates of association strength: single cues. Estimates by control subjects (open bars) and amnesic patients (shaded bars) of the percentage of time that each cue was associated with sunny or rainy weather. Half the subjects were asked these four questions with respect to sunny weather, and half were questioned with respect to rainy weather. Cue 1 was always the cue most strongly associated with the weather outcome in question, and Cue 4 was always the cue least associated with that outcome. Dashed lines indicate the percentage of time that each cue was in fact followed by sunshine or rain (86%, 60%, 40%, and 14%). Error bars indicate the standard error of the mean.

Transfer Tests: Estimates of Association Strength

Single cues. Figure 3 shows the mean estimates of the percentage of time that each cue was followed by sunshine (or rain) when each cue was presented individually. As described above, half the subjects were asked all the questions with respect to sunshine, and half were asked all the questions with respect to rain. The leftmost pair of bars shows the estimates for the cue most strongly associated with whichever weather outcome that a given subject was asked about. The actual percentages are shown by dashed lines (86%, 60%, 40%, and 14%). A 2 (groups) × 4 (cues) ANOVA found no effect of group, $F(1, 22) < 1.0$; an effect of cue, $F(1, 22) = 15.61$, $p < .001$; and a significant interaction of Group × Cue, $F(1, 22) = 5.53$, $p < .03$. The effect of cue reflects the fact that the overall estimates varied in accordance with the different association strengths between each cue and the weather outcome. The interaction between cue and group reflects the fact that the estimates given by the amnesic patients across the cues were less accurate than the estimates given by the control subjects. Specifically, the estimates given by the amnesic patients were not affected by the different association strengths between cue and weather outcome, $F(1, 7) = 1.04$, $p > .30$, but the estimates given by control subjects varied from cue to cue and matched the association strengths rather well, $F(1, 15) = 28.8$, $p < .01$.

Cue pairs. Figure 4 shows the mean estimates of the percentage of time that each cue pair was followed by sunshine

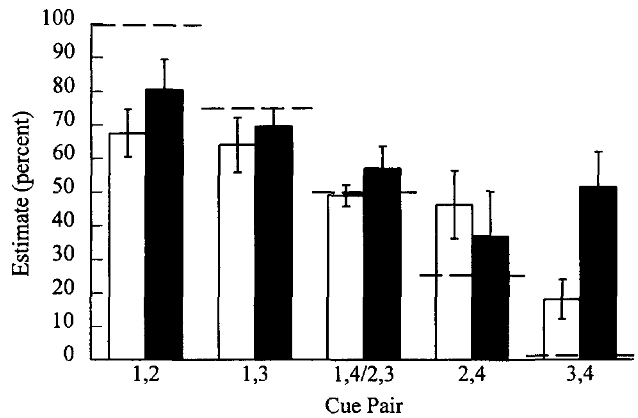

Figure 4. Estimates of association strength: cue pairs. Estimates by control subjects (open bars) and amnesic patients (shaded bars) of the percentage of time that each cue pair was associated with sunny (or rainy) weather. The numbers below the bars identify the two cards in each pair. Half the subjects were asked all the questions with respect to sunny weather, and half were questioned with respect to rainy weather. Cue 1 was always the cue most associated with sunshine (or rain), and Cue 4 was always the cue least associated with sunshine (or rain). The results for Cue Pairs 1, 4 and 2, 3, for which the two outcomes were equally likely, have been averaged together. Dashed lines indicate the percentage of time that each cue pair was in fact followed by sunshine (or rain; 100%, 75%, 50%, 25%, 0%). Error bars indicate the standard error of the mean.

(or rain). Again, half the subjects were asked all the questions with respect to sunshine, and half were asked all the questions with respect to rain. The leftmost pair of bars shows the estimates for the two cues most strongly associated with whichever weather outcome that a given subject was asked about. The actual percentages are shown by dashed lines (100%, 75%, 50%, 25%, and 0%). A 2 (groups) × 5 (cue pairs) ANOVA found a marginal effect of group, $F(1, 22) = 4.19$, $p < .06$; an effect of cue pair, $F(1, 22) = 28.7, p < .0001$; and no interaction, $F(1, 22) < 1.0$. The marginal group effect reflects the tendency for patients with amnesia to give higher estimates than the control subjects. The effect of cue pair reflects the fact that the overall estimates varied in accordance with the different association strengths between cue pairs and the weather outcome. This effect was apparent within each group; for the amnesic patients, $F(1, 7) = 7.44, p < .03$; for the control subjects, $F(1, 15) = 28.7, p < .01$. The two groups did differ in one respect. The control subjects gave a more accurate estimate of association strength than the amnesic patients for the cue pair 3, 4, which was the cue pair most weakly associated with the outcome in question, $F(1, 22) = 8.95, p < .01$.

Cue selection. Performance was assessed with a 1-to-4 scoring system, with 1 being the best score. When one cue was to be selected, subjects were assigned a score of 1 if they correctly chose the cue most strongly associated with each weather outcome, a score of 4 if they chose the least strongly associated cue, and scores of 2 or 3 for the intermediate cues. When two cues were to be selected, subjects were assigned a score based on the ranks of the two cards selected. The ranks of the two scores were summed, and the resulting score was rescaled from a 3-to-7 range to a 1-to-4 range ([0.75 × sum] − 1.25). When three cues were to be selected, a score of 1 was given if the three correct cues were chosen, a score of 4 was given if the most strongly associated cue was not chosen, and a score of 2 or 3 was given for choosing the intermediate cues. In each case, the scores obtained for the two weather outcomes were averaged. Chance performance would result in a score of 2.5.

Figure 5 shows the performance of the amnesic patients and control subjects. Control subjects performed above chance in all conditions, whether they were selecting one, two, or three cues, $ts(15) > 2.64, ps < .03$. In contrast, the amnesic patients performed above chance only when they were asked to select a single cue, $t(7) = 3.27, p < .02$; for two and three cues, their performance was not different from chance ($ts < 1.0$). Finally, the control subjects performed better than the amnesic patients when two cues were to be selected, $F(1, 22) = 14.3, p < .01$.

Questionnaire. Control subjects averaged 6.2 (±0.3) correct responses out of the eight questions, 78% correct (chance was 2.0, or 25% correct). Amnesic patients performed more poorly than control subjects, 4.5 (±0.6) correct answers out of the eight questions, 56% correct, $t(22) = 2.96, p < .01$.

Discussion

Both amnesic patients and control subjects acquired the weather prediction task at a similar rate during 50 trials of training and achieved the same level of performance at the end of training. This finding confirms previous observations that involved three different versions of this same task (Knowlton et al., 1994). The fact that amnesic patients performed as well as control subjects provides strong evidence that performance does not depend materially on declarative memory, which is impaired in amnesia. Despite their normal ability to acquire the prediction task, the amnesic patients performed poorly on the eight-item questionnaire, which asked simple questions about the training episode. Overall, the amnesic patients also

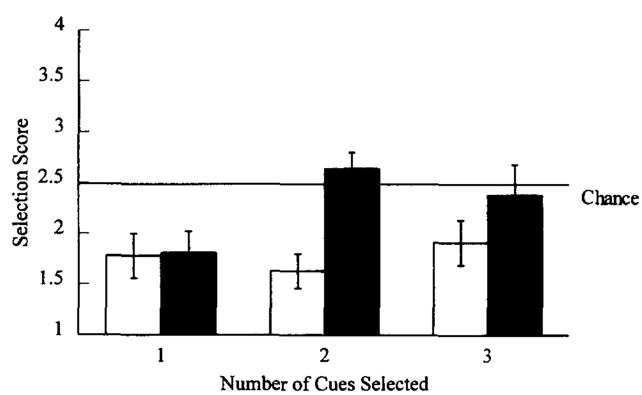

Figure 5. Cue selection. Selection scores obtained by control subjects (open bars) and amnesic patients (shaded bars) when they chose the one cue (or the two cues, or the three cues) most likely to be showing if the weather outcome was sunshine (or rain). A lower selection score indicates better performance. Chance performance (indicated by horizontal line) is 2.5 for all three tasks. Error bars indicate the standard error of the mean.

performed more poorly than control subjects on three transfer tests involving a total of 16 questions that asked about their acquired knowledge, specifically, questions that asked them to estimate the association strength between the four cues and the weather outcome. Nevertheless, the amnesic patients did perform as well as control subjects on some questions. They were able to estimate the association strength of cue pairs about as well as control subjects (see Figure 4) (except for the most weakly associated cue pair), and they were able to select the single cue most strongly associated with sunshine or rain (see Figure 5).

It is possible that amnesic patients were sometimes able to perform satisfactorily because certain of the questions essentially reinstantiated the task conditions and did not require flexible knowledge of the cue–outcome associations. By contrast, the patients may have performed poorly on many other questions because these questions did require flexible use of their task knowledge. We explored this possibility in Experiment 2 by obtaining ratings of each question from an independent group of control subjects. We asked the subjects to rank the questions according to whether they required more or less flexible use of knowledge. We expected that the estimates requiring the most flexibility should be the ones on which the amnesic patients were impaired.

Experiment 2

The purpose of this experiment was to obtain independent estimates of the extent to which the questions used in Experiment 1 required flexible use of knowledge about the association between cues and weather outcomes.

Method

Participants

The subjects were 2 men and 18 women (mean age = 25.8 years of age, mean years of education = 16.0) from the UCSD and VA communities.

Materials and Procedure

Subjects completed 50 trials of training on the weather prediction task exactly as in Experiment 1 followed by the same three transfer tests from Experiment 1. In addition, after completing each of the three estimation tasks, subjects were asked to rank the questions in terms of how directly or indirectly the questions in each task assessed the knowledge acquired during the weather prediction task. It was explained that "A direct question simply asks you to make the same kind of decision you have already been making. An indirect question asks you to use what you learned in a new and flexible way," and "the direct questions tend to put you in the same situation you were in when you were predicting the weather from the cards while the indirect questions create a different situation in which you have to apply what you learned."

In this way, the objective was to obtain judgments concerning which questions simply reinstantiated the task conditions and which questions required that subjects use what they had learned in a flexible way. For the estimates of single cues, subjects ranked four questions (1 = *most direct,* 4 = *least direct*). Subjects indicated their rankings by ordering the four cues, which were depicted on separate index cards. For the estimates of cue pairs, subjects ranked six questions (1 = *most direct,* 6 = *least direct*) using six cards, each showing one of the possible cue pairs. For the cue selection test, subjects placed the six questions into three groups; that is, they ranked the questions that involved selecting one, two, or three cards (disregarding the two weather outcomes; 1 = *most direct,* 3 = *least direct*).

Results

Weather Prediction

This group obtained scores of 61%, 68%, 62%, 63%, and 66% correct during five training blocks (10 trials/block). Although performance did not measurably improve across the five blocks, $F(1, 19) = 0.20$, this group did not perform differently from the control group in Experiment 1. A 2 (groups) × 5 (blocks) ANOVA found no effect of group, $F(1, 34) = 0.16$; a marginal trend for performance to increase across blocks, $F(1, 34) = 3.75, p < .07$; and no interaction, $F(1, 34) = 2.09$, $p > .15$. Moreover, the two groups reached a similar level of performance on the final block of 10 trials (66% vs. 68%), and both scores were well above chance levels (ps $< .01$).

Transfer Tests: Estimates of Association Strength

All of the estimates resembled the estimates given by control subjects in Experiment 1.

Single cues. Subjects gave mean cue–outcome association estimates of 74%, 52%, 58%, 20% (compare with Figure 3). A 2 × 4 ANOVA comparing this group to the control subjects in Experiment 1 indicated no effect of group $F(1, 34) < 1.0$; a significant effect of cue, $F(1, 34) = 77.7, p < .0001$; and no interaction, $F(1, 34) < 1.0$.

Cue pairs. Subjects gave mean cue–pair association estimates of 65%, 69%, 52%, 28%, and 30% (compare with Figure 4). A 2 × 5 ANOVA comparing this group to the control subjects in Experiment 1 (similar to above) indicated no effect of group, $F(1, 34) < 1.0$; a significant effect of cue pair, $F(1, 34) = 73.2, p < .0001$; and no interaction, $F(1, 34) < 1.0$.

Cue selection. Subjects scored 1.60, 1.79, and 1.83 when selecting one, two, and three cues, respectively (compare with Figure 5). They did not perform differently than the control subjects on any of the three tests (Fs < 1.0).

Ranking Questions as Direct or Indirect

Single cues. Figure 6A shows the average ranks given to the questions about each cue. Lower scores indicate that the question was judged to be relatively direct, and higher scores indicate that the question was judged to require more indirect, flexible use of task knowledge. The scores increased as the association between the cue and weather outcome decreased, $F(1, 19) = 12.97, p < .01$. Thus, subjects indicated that making association-strength estimates for cues weakly associated with a weather outcome required more flexibility than making estimates for strongly associated cues.

Cue pairs. Figure 6B shows the average ranks given to the questions about each cue pair. The scores tended to increase as the association between the cue and weather outcome decreased, $F(1, 19) = 3.77, p < .07$. Thus, just as with single

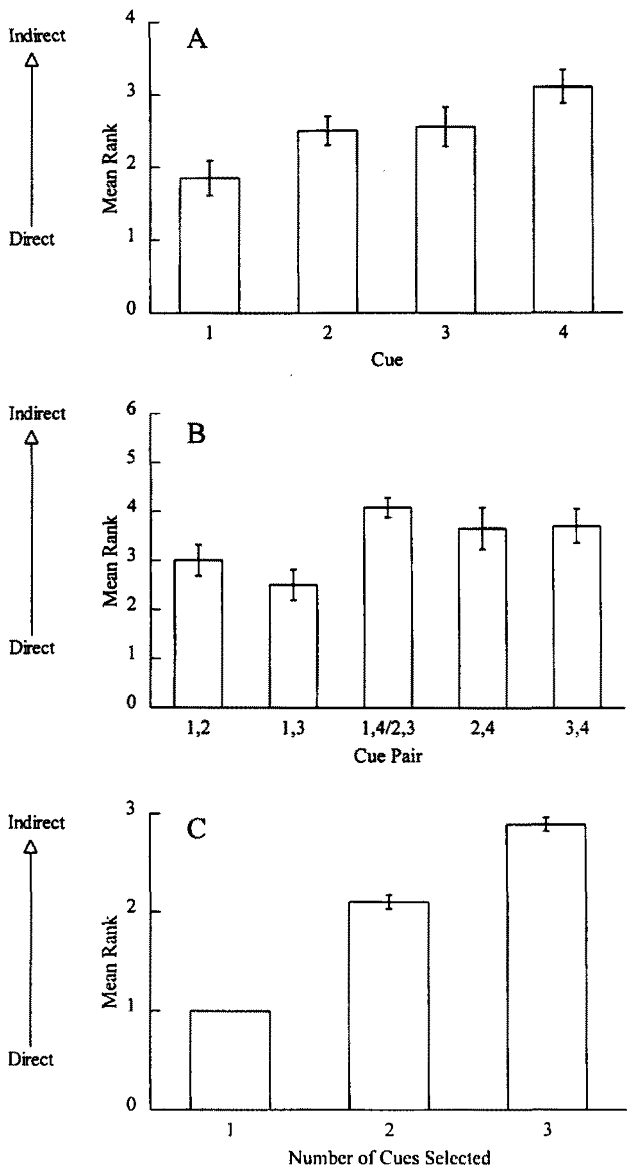

Figure 6. Ranking questions as direct or indirect. A higher rank identifies questions judged as more indirect and therefore one requiring relatively more flexible use of task knowledge. Cue 1 was always the cue most associated with the weather outcome in question (sunshine or rain), and Cue 4 was always the cue least associated with that outcome. Error bars indicate the standard error of the mean. A: Rankings for the four questions about the association strength of single cues. B: Rankings for the questions about the association strength of single cues. The two cue pairs (1, 4 and 2, 3) for which both outcomes are equally predicted are averaged. C: Rankings for the three questions that asked subjects to select cues given a particular weather outcome.

cues, subjects indicated that making association-strength estimates for weakly associated cue pairs required flexible knowledge.

Numerically, the questions about the middle two cue pairs (1, 4 and 2, 3) were judged to be the most indirect (center bar, Figure 6B). Because these two cue pairs were associated equally often with both outcomes and because each of these pairs was presented only twice (as opposed to four times for the other pairs), it is unclear how to interpret the ranking given these two pairs. For example, these pairs were presented so infrequently that subjects may have tried to infer their association strength indirectly from information about the individual cues.

Cue selection. Figure 6C shows the average rank given to each question type. Scores increased as subjects were asked to select an increasing number of cues, $F(1, 18) = 686, p < .001$. That is, subjects judged the task of selecting one cue as the most direct use of their knowledge, and they indicated that selecting two or three cues involved progressively more indirect and flexible use of their task knowledge. Indeed, 17 of the 20 subjects ranked the questions in this order.

Discussion

To make association strength estimates for both single cues (Figure 6A) and cue pairs (Figure 6B), subjects indicated that more flexible use of task knowledge was required as the association between the cue (or cue pair) and the weather outcome became weaker. That is, subjects reported that flexible use of task knowledge was needed when they had to estimate the association strength of weak associations between cues and the weather outcome. This result probably occurred because a weak cue–outcome association implies that the cue (or cue pair) in question had been frequently associated with the opposite weather outcome. To give an estimate of association strength, subjects may need to draw on this knowledge (that the cue or cue pair in question is strongly associated with the opposite weather outcome), and this strategy requires flexible use of their knowledge. For the questions that asked subjects to select a cue or cues (cue selection), subjects indicated that selecting one card required the least flexible use of task knowledge and that selecting two or three cards required more flexible use of task knowledge (see Figure 5).

General Discussion

In Experiment 1, amnesic patients exhibited normal learning of the probabilistic classification task (weather prediction) across 50 training trials. The flexibility of their task knowledge was then assessed with transfer tests that asked them to estimate in various ways the association strength between the four cues and the two weather outcomes. The amnesic patients were often impaired in comparison with control subjects at estimating the association strengths of cues (see Figures 3 and 4), and they were also impaired at selecting cues most strongly associated with each weather outcome (see Figure 5). In Experiment 2, an independent group of control subjects judged to what extent the transfer test simply recreated the conditions of the learning task or created new conditions that required flexible use of task knowledge. In general, amnesic patients failed the same questions that control subjects rated as assessing task knowledge indirectly and requiring flexible use of task knowledge.

This relationship was particularly clear for the cue selection test (see Figures 5 and 6C). Figure 6C shows that the questions asking subjects to select either two or three cues for a specific weather outcome were overwhelmingly rated as requiring more flexibility than the questions that asked for selection of only a single cue. Figure 5 shows the performance of amnesic patients and control subjects on these questions. The patients performed at chance levels when two or three cues were to be selected but performed similarly to control subjects when one cue was to be selected (see Experiment 1, Results). The relationship between test performance and ratings of flexibility can also be observed for the test in which subjects estimated the percentage of time that each cue pair was associated with sunny (or rainy) weather (see Figure 4). Figure 4 shows that amnesic patients were most impaired at estimating the association strength of cue pair 3, 4. Figure 6 shows that estimating the association strength of this cue pair was rated as requiring more flexible use of knowledge than most of the other cue pairs.

The amnesic patients were not impaired at providing estimates of association strength for the two cue pairs 1, 4 and 2, 3 (Figure 4, center bar), although these cue pairs were rated as requiring the most flexibility (Figure 6B, center bar). However, for these two questions, it is unclear how to interpret the estimates given by the patients. If patients had been unable to make estimates for these cue pairs, they might have chosen a value in the middle of the 0%–100% response scale (i.e., 50%, which was in fact the correct response for these two cue pairs). Note that in the other cases when amnesic patients performed poorly (see Figure 3 and Figure 4, cue pair 3, 4) they did choose values close to 50%, but in those cases 50% was incorrect.

Note also that the amnesic patients performed well on transfer test questions that were rated as assessing task knowledge rather directly (i.e., questions that did not require flexible use of task knowledge). Figure 6 identifies four questions that were rated as assessing task knowledge most directly: estimating the association strength of a single cue (Figure 6A), estimating the association strength of cue pairs 1, 2 and 1, 3 (Figure 6B), and selecting the single cue most associated with a specific weather outcome (Figure 6C). The performance of amnesic patients on these four questions was nearly as good as the performance of control subjects. Amnesic patients gave estimates of 65%, 81%, and 70% for the first three of these questions and obtained a selection score of 1.8 on the fourth question. Control subjects gave estimates of 80%, 68%, and 64%, respectively, and obtained a selection score of 1.8. None of these pairwise comparisons approached significance, $ts(22) < 1.35$, $ps > .20$. The finding that the amnesic patients performed well on transfer tests that assessed their acquired knowledge rather directly provides additional evidence that the patients learned the classification task as well as control subjects during the 50 training trials. The findings also highlight the difficulty that the patients had with the more indirect tests of their task knowledge.

The finding that amnesic patients acquired the probabilistic classification task (weather prediction) at a normal rate indicates that this learning occurred independently of the brain structures damaged in amnesia. Performance by amnesic patients on this task presumably did not depend materially on declarative memory because declarative memory is severely impaired in amnesia. Furthermore, the finding that amnesic patients could not perform well on the transfer tests presumably reflects the characteristics of nondeclarative memory; namely, nondeclarative knowledge about how to perform the classification task itself cannot be applied flexibly to novel situations. Unlike the amnesic patients, control subjects had the opportunity to acquire declarative knowledge about the task in parallel to learning how to respond appropriately to its probabilistic structure. Thus, unlike the amnesic patients, control subjects were able to answer questions about the events of training (the eight-item questionnaire), and they were able to transfer their knowledge about cue–outcome relationships to tasks that asked indirect questions about the association strengths between the cues and outcomes. The superior performance of the control subjects on the transfer tasks is apparently due to concomitant declarative knowledge accrued during the learning of the classification task. This knowledge did not contribute to weather prediction performance during the 50 training trials, but it was available for the transfer test to support estimates of association strengths between the cues and the two weather outcomes.

Note that the results cannot be explained by supposing that the amnesic patients acquired the prediction task within short-term working memory (which is intact in amnesia) but then forgot what they had learned by the time they were given the transfer test. In our previous study (Knowlton et al., 1994), amnesic patients learned this same task during 50 training trials and then continued to perform well after a 5-min delay filled with conversation.

It should be noted that several paradigms for studying nondeclarative memory in patients with amnesia rely on some ability to transfer knowledge acquired during a training episode to novel stimuli (e.g., artificial grammar learning, Knowlton et al., 1992; prototype learning, Knowlton & Squire, 1993) that are structurally similar but not identical to training items. However, the type of flexibility demanded by the test used here requires that the acquired information be used in a distinctly new way. Our results suggest that declarative and nondeclarative forms of memory differ when considerable flexibility is required; that is, the ability to use nondeclarative memory flexibly is limited in comparison with the ability to use declarative memory flexibly.

Much of the available evidence about declarative and nondeclarative memory systems has come from lesion studies and from task dissociations in normal subjects. It is now clear that these memory systems can also be distinguished by their operating characteristics, as one might expect if different brain systems are involved in different kinds of memory. To dissociate memory into functionally separate systems, however, it is important to determine how operating characteristics differ across the systems. For example, the systems differ with respect to their capacity to support awareness of memory content. Declarative memory supports conscious recollections about facts and events, whereas nondeclarative memory is expressed through performance without awareness of memory content or even that memory is being tested. In addition, declarative memory is flexible (i.e., accessible across broad testing conditions), whereas nondeclarative memory is less flexible or "hyperspecific" and is best accessed in conditions similar to the circumstances or original learning.

This idea about declarative and nondeclarative memory has been proposed previously (Cohen, 1984; Tulving & Schacter, 1990) and has been supported by several studies with experimental animals and humans (Eichenbaum et al., 1989, 1990; Glisky et al., 1986a, 1986b; Saunders & Weiskrantz, 1989). However, most of these studies (Eichenbaum et al., 1989, 1990; Glisky et al., 1986a, 1986b) involved tasks in which the rate of initial learning was abnormally slow. Accordingly, it was possible in those studies that the inflexibility observed in transfer tests was due to the inflexibility of whatever declarative memory could be gradually acquired during extended training. In the present study, however, amnesic patients learned at a normal rate. Therefore, learning was presumably supported by nondeclarative memory. The impaired performance of the amnesic patients on the transfer tests indicates that inflexibility is a property of nondeclarative memory, not a property of residual declarative memory that is acquired at an abnormally slow rate.

Although we used subjective ratings to provide an indication of which transfer tests required more flexible use of task knowledge, the approach here is nevertheless similar to studies of memory flexibility in experimental animals (e.g., Eichenbaum et al., 1989). In both cases, one first identifies a task that depends primarily on nondeclarative memory. Transfer tests are then constructed that require task knowledge to be applied flexibly. Performance on the transfer tests is then compared for subjects with either intact or impaired declarative memory. An additional method available only in human studies is to collect subjective ratings to further refine the analysis of the transfer test results.

The finding that declarative and nondeclarative memory systems differ with respect to the flexibility of the knowledge acquired by each system provides a useful and important step in understanding the organization of memory. First, unlike differences in awareness, differences in the flexibility of memory can be studied in both humans and experimental animals. Second, information about the operating characteristics of memory systems should inform neurobiological studies directed at the anatomy and physiology of related brain systems. In the case of the distinction between declarative and nondeclarative memory relevant to the probabilistic classification task, current information suggests that the focus will be on the medial temporal lobe and neocortex, on the one hand (Squire, 1992), and the neocortex and neostriatum (Knowlton, Mangels, & Squire, in press; Knowlton, Paulsen, et al., in press; Knowlton et al., 1994; Mishkin & Petri, 1984), on the other.

References

Butters, M., Glisky, E. L., & Schacter, D. L. (1993). Transfer of new learning in memory-impaired patients. *Journal of Clinical and Experimental Neuropsychology, 15,* 219–230.

Cohen, N. J. (1984). Preserved learning capacity in amnesia: Evidence for multiple memory systems. In L. R. Squire & N. Butters (Eds.), *Neuropsychology of memory* (pp. 83–103). New York: Guilford Press.

Eichenbaum, H., Mathews, P., & Cohen, N. J. (1989). Further studies of hippocampal representation during odor discrimination learning. *Behavioral Neuroscience, 103,* 1207–1216.

Eichenbaum, H., Stewart, C., & Morris, R. G. M. (1990). Hippocampal representation in place learning. *Journal of Neuroscience, 10,* 3531–3542.

Gilbert, J., Levee, R., & Catalano, K. (1968). A preliminary report on a new memory scale. *Perceptual and Motor Skills, 27,* 277–278.

Glisky, E. L., Schacter, D. L., & Tulving, E. (1986a). Computer learning by memory-impaired patients: Acquisition and retention of complex knowledge. *Neuropsychologia, 24,* 313–328.

Glisky, E. L., Schacter, D. L., & Tulving, E. (1986b). Learning and retention of computer-related vocabulary in memory-impaired patients: Method of vanishing cues. *Journal of Clinical and Experimental Neuropsychology, 8,* 292–312.

Janowsky, J. S., Shimamura, A. P., Kritchevsky, M., & Squire, L. R. (1989). Cognitive impairment following frontal lobe damage and its relevance to human amnesia. *Behavioral Neuroscience, 103,* 548–560.

Kaplan, E. F., Goodglass, H., & Weintraub, S. (1983). *The Boston Naming Test.* Philadelphia: Lea Febiger.

Knowlton, B. J., Mangels, J. A., & Squire, L. R. (in press). A neostriatal habit learning system in humans. *Science.*

Knowlton, B. J., Paulsen, J. S., Squire, L. R., Swerdlow, N., Swenson, M., & Butters, N. (in press). Dissociations within nondeclarative memory in Huntington's disease. *Neuropsychology.*

Knowlton, B. J., Ramus, S. J., & Squire, L. R. (1992). Intact artificial grammar learning in amnesia: Dissociation of classification learning and explicit memory for specific instances. *Psychological Science, 3,* 172–179.

Knowlton, B. J., & Squire, L. R. (1993). The learning of natural categories: Parallel memory systems for item memory and category-level knowledge. *Science, 262,* 1747–1749.

Knowlton, B. J., Squire, L. R., & Gluck, M. (1994). Probabilistic classification learning in amnesia. *Learning and Memory, 1,* 106–120.

Kritchevsky, M., Squire, L. R., & Zouzounis, J. A. (1988). Transient global amnesia: Characterization of anterograde and retrograde amnesia. *Neurology, 38,* 213–219.

Mattis, S. (1976). Dementia Rating Scale. In R. Bellack & B. Keraso (Eds.), *Geriatric psychiatry* (pp. 77–121). New York: Grune & Stratton.

Mishkin, M., & Petri, H. L. (1984). Memories and habits: Some implications for the analysis of learning and retention. In L. R. Squire & N. Butters (Eds.), *Neuropsychology of memory* (pp. 287–296). New York: Guilford Press.

Nadel, L. (1994). Multiple memory systems: What and why, an update. In D. L. Schacter & E. Tulving (Eds.), *Memory systems 1994* (pp. 39–64). Cambridge, MA: MIT Press.

Nissen, M. J., & Bullemer, P. (1987). Attentional requirements of learning: Evidence from performance measures. *Cognitive Psychology, 19,* 1–32.

Osterrieth, P. A. (1944). Le test de copie d'une figure complexe [The test of copying a complex figure]. *Archives de Psychologie, 30,* 206–356.

Polich, J., & Squire, L. R. (1993). P300 from amnesic patients with bilateral hippocampal lesions. *EEG Clinical Neurophysiology, 86,* 408–417.

Reber, A. S. (1989). Implicit learning and tacit knowledge. *Journal of Experimental Psychology: General, 118,* 219–235.

Reber, P. J., & Squire, L. R. (1994). Parallel brain systems for learning with and without awareness. *Learning and Memory, 2,* 1–13.

Roediger, H. (1990). Implicit memory: Retention without remembering. *American Psychologist, 45,* 1043–1056.

Saunders, R. C., & Weiskrantz, L. (1989). The effects of fornix transection and combined fornix transection, mammillary body lesions and hippocampal ablations on object–pair association memory in the rhesus monkey. *Behavioral Brain Research, 35,* 85–94.

Schacter, D. L. (1992). Understanding implicit memory: A cognitive neuroscience approach. *American Psychologist, 47,* 559–569.

Schacter, D. L., & Tulving, E. (1994a). *Memory systems 1994.* Cambridge, MA: MIT Press.

Schacter, D. L., & Tulving, E. (1994b). What are the memory systems of 1994? In D. L. Schacter & E. Tulving (Eds.), *Memory systems 1994* (pp. 1–38). Cambridge, MA: MIT Press.

Shanks, D. R., & St. John, M. F. (1994). Characteristics of dissociable human learning systems. *Behavioral and Brain Sciences, 17,* 367–447.

Sherry, D. F., & Schacter, D. L. (1987). The evolution of multiple memory systems. *Psychological Review, 94,* 439–454.

Shimamura, A. P., & Squire, L. R. (1988). Long-term memory in amnesia: Cued recall, recognition memory, and confidence ratings. *Journal of Experimental Psychology: Learning, Memory and Cognition, 14,* 763–770.

Squire, L. R. (1992). Memory and the hippocampus: A synthesis from findings with rats, monkeys, and humans. *Psychological Review, 99,* 195–231.

Squire, L. R. (1994). Declarative and nondeclarative memory: Multiple brain systems support learning and memory. In D. Schacter & E. Tulving (Eds.), *Memory systems 1994.* Cambridge, MA: MIT Press.

Squire, L. R., Amaral, D. G., & Press, G. A. (1990). Magnetic resonance measurements of hippocampal formation and mammillary nuclei distinguish medial temporal lobe and diencephalic amnesia. *Journal of Neuroscience, 10,* 3106–3117.

Squire, L. R., & Shimamura, A. P. (1986). Characterizing amnesic patients for neurobehavioral study. *Behavioral Neuroscience, 100,* 866–877.

Squire, L. R., & Zola-Morgan, S. (1991). The medial temporal lobe memory system. *Science, 253,* 1380–1386.

Tulving, E. (1991). Concepts in human memory. In L. R. Squire, N. M. Weinberger, G. Lynch, & J. L. McGaugh (Eds.), *Memory: Organization and locus of change* (pp. 3–32). New York: Oxford University Press.

Tulving, E., & Schacter, D. L. (1990). Priming and human memory systems. *Science, 247,* 301–306.

Warrington, E. K. (1984). *Recognition Memory Test.* Windsor, England: FER-Nelson.

Weiskrantz, L. (1991). Problems of learning and memory: One or multiple memory systems? In J. R. Krebs & G. Horn (Eds.), *Behavioural and neural aspects of learning and memory* (pp. 1–10). Oxford, England: Clarendon Press.

Willingham, D. B., Nissen, M. J., & Bullemer, P. (1989). On the development of procedural knowledge. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 15,* 1047–1060.

Received July 25, 1995
Revision received October 24, 1995
Accepted January 12, 1996 ■