Journal of the International Neuropsychological Society (2003), **9**, 394–406.

DOI: 10.1017/S1355617703930050

Categorization and recognition performance of a memory-impaired group: Evidence for single-system models

SAFA R. ZAKI,[1] ROBERT M. NOSOFSKY,[1] NENETTE M. JESSUP,[2] AND FREDERICK W. UNVERZAGT[2]
[1]Department of Psychology, Indiana University, Bloomington
[2]Department of Psychiatry, Indiana University School of Medicine, Indianapolis

(Received January 7, 2002; Revised April 16, 2002; Accepted April 30, 2002)

Abstract

Previous research has demonstrated dissociations between categorization and recognition performance in amnesic patients, supporting the idea that separate memory systems govern these tasks. However, previous research has also demonstrated that these dissociations are predicted by a single-system model that allows for reasonable parameter differences across groups. Generally, previous studies have employed categorization tasks that are less demanding than the recognition tasks. In this study, we distinguish between single-system and multiple-system accounts by testing memory-impaired individuals in a more demanding categorization task. These patients, just like previous amnesic participants, show a dissociation between categorization and recognition when tested in previously employed paradigms. However, they display a categorization deficit when tested in the more challenging categorization task. The results are interpreted as support for a single-system framework in which categorization and recognition depend on one representational system. (*JINS*, 2003, *9*, 394–406.)

Keywords: Categorization, Recognition, Memory impairment, Categorization recognition dissociations

INTRODUCTION

Recently, much attention has been given to the issue of whether a single memory system underlies categorization and old-new recognition. Various dissociations between categorization and recognition have been reported in support of a multiple-systems framework (Knowlton & Squire, 1993; Knowlton et al., 1994; Reber & Squire, 1999; Reed et al., 1999; Squire & Knowlton, 1995). Generally, in these studies, amnesics classify stimuli into various categories just as well as do normal controls, but perform significantly worse on tests of explicit memory, such as recognition and recall.

For example, in an intriguing and widely cited study, Knowlton and Squire (1993) tested amnesics and normal controls in categorization and recognition tasks involving the classic Posner and Keele (1968, 1970) dot-pattern paradigm. In the categorization task, participants observed 40 high-level distortions of a dot-pattern prototype and were subsequently tested with various new category members and nonmembers. In a recognition task, participants viewed five random patterns eight times each and then were given an old/new test involving the five old and five new patterns. Knowlton and Squire's (1993) key result was that although the amnesics performed at significantly reduced levels in the recognition task, their performance on the categorization task was at statistically normal levels. Knowlton and Squire (1993) interpreted this finding as providing support for a multiple-systems approach. According to their interpretation, an explicit memory system, which is damaged in amnesics, is responsible for old–new recognition; whereas a procedural system, which is intact in amnesics, is responsible for categorization.

Similarly, Reed et al. (1999) generalized the Knowlton and Squire (1993) findings to object-like stimuli. Participants in their study trained on line drawings of cartoon animals that varied on nine binary-valued dimensions. In a training phase, participants were exposed to a category of these animals by viewing various low-level distortions of a

Reprint requests to: Safa Zaki, Department of Psychology, Brontman Science Center, Williams College, 18 Hoxsey Street, Williamstown, MA 01267. E-mail: szaki@williams.edu

prototype. In a test that followed, they were asked to indicate whether or not particular new items were members of the previously viewed category. The test items included the prototype, some new low-level distortions, some neutral items, and various non-members of the category. In addition, the participants were given an explicit memory test in which they were asked to recall the two possible values along each of the nine dimensions. Reed et al. (1999) found that although the amnesic patients were significantly impaired at the recall task, they were able to categorize the cartoon animals at levels within the normal range. Once again, these findings were taken as support for a multiple-systems account of recognition and categorization.

However, both these key results have been shown to also be congruent with a single-system account of categorization and recognition. For example, Nosofsky and Zaki (1998) proposed a single-system exemplar-model account of Knowlton and Squire's (1993) data. Specifically, according to the generalized context model (Nosofsky, 1986), there is only one representational system in which exemplars are stored, and this representational system is accessed when making both categorization and recognition decisions (for similar ideas, see Heit, 1993; Hintzman, 1986; 1988; Medin & Shaffer, 1978). Furthermore, according to this approach, the ability to discriminate between items in memory is impaired in amnesia. In the model, this ability is represented in terms of a sensitivity parameter. The basic tenet behind this approach was that, in the tasks used by Knowlton and Squire (1993), adequate categorization performance requires only a low level of sensitivity, whereas recognition requires higher sensitivity levels. (For related ideas advanced to explain dissociations between categorization and recognition in the domain of artificial grammar learning, see Kinder & Shanks, 2001.)

Nosofsky and Zaki (1998) provided two lines of evidence in support of this single-system hypothesis. First, they induced lower sensitivity levels in college-aged normal participants by introducing a delay between the training and testing session of Knowlton and Squire's (1993) paradigm. Participants in this experiment completed the recognition or categorization test phase either immediately after exposure to the training stimuli or after a 1-week delay. The results of this experiment closely resembled those obtained by Knowlton and Squire (1993). That is, just like amnesia, delay caused a small decrement in classification, and a substantial decrement in recognition. Thus, a general memory impairment is more detrimental to old-new recognition than categorization in Knowlton and Squire's (1993) paradigm.

In a second approach, Nosofsky and Zaki (1998) conducted formal modeling of Knowlton and Squire's (1993) data. In this analysis, they demonstrated that a lower setting of the sensitivity parameter in the exemplar model predicts a large drop in recognition performance but only a slight drop in categorization performance. Indeed, a simple version of the model provided a good quantitative fit to the complete set of categorization and recognition data reported by Knowlton and Squire (1993). In sum, both the empirical and model-based approaches used by Nosofsky and Zaki (1998) supported the idea that different task requirements in categorization and recognition might be the underlying reason for the observed dissociation.

Likewise, Zaki and Nosofsky (2001) demonstrated that the Reed et al. (1999) findings involving object-like stimuli were consistent with a single-system account. Zaki and Nosofsky (2001) noted that in the categorization paradigm used by Reed et al. (1999), the values of the dimensions within the category were highly correlated such that a participant could focus attention on any single dimension and still achieve impressive accuracy levels. However, no such strategy was available in the explicit memory task (cued-recall or recognition), which required attention to many of the dimensions of the objects in order to achieve adequate performance. Therefore, Zaki and Nosofsky (2001) hypothesized that this difference in task requirements might be sufficient to explain the observed dissociation.

To investigate these ideas, Zaki and Nosofsky (2001) tested normal participants in the Reed et al. (1999) task. The main change introduced in Zaki and Nosofsky's (2001) version of the task was that for some participants the test phase was delayed by 1 week. Their results indicated that although participants showed a dramatic decrease in old–new recognition performance after a 1-week delay, they were not significantly worse in the categorization task. More importantly, Zaki and Nosofsky (2001) demonstrated through formal modeling of the data that subjects in the categorization task were indeed paying attention to only a small subset of the dimensions that composed the objects. By contrast, good performance in the recognition task was dependant on using information from many of the dimensions. Thus, once again, the dissociation taken as evidence of the existence of multiple systems could be simply explained in terms of differences in the memory demands between the categorization and recognition tasks.

Even data from the severely amnesic patient E.P. (Squire & Knowlton, 1995; Squire & Zola, 1996), once thought to be severely damaging to a single-system account of categorization and recognition, can now potentially be explained within this framework. In the dot-pattern paradigm, Squire and Knowlton (1995) demonstrated that E.P. literally showed zero discrimination between old and new exemplars in the recognition task yet performed normally on the categorization task. Similarly, Reed et al. (1999) demonstrated that E.P. could not recall any of the dimensions of the stimuli, but performed at the same levels as controls in the categorization task. However, Palmeri and Flanery (1999) and Zaki and Nosofsky (2001) showed that subjects can perform these particular categorization tasks without any prior exposure to the training set, presumably by making use of the similarity of items in the test to infer the category. Therefore, in both these tasks, E.P. may have used his available working memory to complete the categorization task without having any long-term memory for the original training instances.

In summary, both single-system and multiple-system approaches appear to be compatible with the categorization–recognition dissociations observed by Knowlton and Squire (1993), Squire and Knowlton (1995), and Reed et al. (1999). The purpose of the present research was to begin to make headway toward developing contrasts between the predictions of these approaches. Our key idea was as follows: If categorization and recognition are indeed mediated by separate representational systems, with the categorization system in memory-impaired individuals left intact, then the dissociation between categorization and recognition should continue to be observed even if the categorization task is made more demanding. By contrast, if the hypothesis of a single representational system is correct, then memory-impaired individuals who perform at normal levels on the Knowlton-Squire (1993) categorization task should show clear deficits on a more demanding categorization task. Although the amnesics' lowered level of memory sensitivity may have been sufficient to enable them to perform at near-normal levels on Knowlton and Squire's (1993) task, this memory impairment might be far more detrimental on an alternative categorization task.

There are two aspects of Knowlton and Squire's (1993) and Reed et al.'s (1999) categorization tasks that may have contributed to a lack of diagnosticity. The first is that the observers experienced only a single category of objects during the training phase. By contrast, most standard categorization tasks require individuals to form representations of multiple categories and to distinguish among them. Second, in the test phase of Knowlton and Squire's (1993) and Reed et al.'s (1999) categorization tasks, observers were tested on new objects only. In these paradigms, the researchers did not present any of the original training instances at time of test. The single-system exemplar model predicts, however, that a lowered level of memory sensitivity should be more detrimental to performance on old items than on new items. Therefore, including old training instances as test items should increase the diagnosticity of the categorization experiment.

In the present experiment, we tested patients with memory disorders (described in greater detail below), as well as a group of normal controls matched on age and education, on a battery of categorization and recognition conditions. Among the conditions that we included were the original categorization and recognition tasks used by Knowlton and Squire (1993). Because the patients we tested are known to primarily have deficits on explicit memory tasks, we expected the patients and matched normal controls to show the same dissociation between categorization and recognition as was observed in the original study of Knowlton and Squire (1993).

The key test in the battery was a new two-category paradigm that we believed would make greater demands on observers' memory resources and that might therefore reveal a deficit in the patient group. The two-category experiment involved stimuli that were essentially the same as those used in Knowlton and Squire's (1993) single-category task, that is, dot-pattern distortions generated from prototype patterns. However, rather than learning only a single category of objects, observers now needed to form distinct representations of two separate categories of patterns. In addition, at time of test, rather than presenting only new objects, we tested observers with the original training instances as well. The prediction from the single-system exemplar model is that although the memory-impaired patients would display near-normal performance on the single-category task of Knowlton and Squire (1993), they would display a significant performance deficit on the two-category task. Some previous research has also examined the ability of amnesics to learn two-category structures (Filoteo et al., 2001; Knowlton et al., 1994; Kolodny, 1994), with mixed results. We discuss the relation between the present experiment and this other research in our General Discussion.

We tested two groups of memory-impaired patients in the present experiment. Some of the patients were diagnosed with amnestic disorder while others were diagnosed with dementia due to probable Alzheimer's disease. There is a very thin diagnostic line between these two types of patients in the clinical sense, and we hoped to be able to combine the results of these two groups in order to gain more statistical power. One potential caveat is that although the subjects with amnestic disorder by definition have an isolated memory impairment, other impairments must accompany memory loss in the demented group. Therefore, a potential difficulty of interpretation is that if the demented patients perform worse than the matched normal controls on the two-category task, it could be attributed to their more widespread cognitive impairment. We followed two approaches to addressing this concern. First, we sought evidence that the two patient groups performed similarly on all of the experimental tasks before combining their results. Second, and most critically, we evaluated the extent to which we replicated the categorization–recognition dissociation on the *original* Knowlton and Squire (1993) tasks. According to the dissociation logic advanced by Knowlton and Squire (1993), if the patients perform as well as do the controls on the original single-category task, it provides evidence that their category-learning system has been spared.

METHODS

Participants

Participants were drawn from the clinical and research programs on aging and dementia within the Departments of Psychiatry and Neurology at Indiana University School of Medicine. Two sets of memory-impaired participants were recruited for this project. Seven subjects met *DSM–IV* (American Psychiatric Association, 1994) criteria for amnestic disorder and 10 subjects met NINCDS/ADRDA criteria for dementia due to probable Alzheimer disease (McKhann et al., 1984).[1] Twelve healthy controls were recruited from the Indiana Alzheimer Disease Center registry. A consensus diagnostic conference consisting of neurologists, psychiatrists, neuropsychologists, and nurses found the controls to be functioning in the normal range based on physician examination, neuropsychological testing, and extensive informant interview that reviewed symptoms and adequacy in daily functioning. The control group was matched in terms of age and education to the patient sample. The patients averaged 12.82 years of education and were on average 76.00 years old at the time of testing. The normal controls averaged 14.83 years of education and were on average 73.67 years of age at the time of the study. The two groups did not differ significantly on age $[t(27) = 1.412, p > .10]$ or education $[t(27) = 1.711, p = .10$.

All participants completed a short battery of neuropsychological tests prior to engaging in the experimental paradigm. The following tests were used to characterize the cognitive and affective status of the sample:

1. *Mini-Mental State Examination* (Folstein et al., 1975) is a screening test of general cognition that assesses orientation to time and place, verbal memory, attention, language, and visuoconstructional skill. The attention item was spelling the word *world* backward. Scores range from zero to 30.
2. *Boston Naming Test* (Goodglass & Kaplan, 1983) is a test of confrontation naming. Line drawings of objects of varying degrees of frequency in experience are presented one at a time from a stimulus book. Subjects are to state the name of the object. Circumlocutions and descriptions do not receive credit. Scores range from zero to 60.
3. *Constructional Praxis* (Morris et al., 1988) is a test of graphomotor skill in which subjects copy geometric figures (circle, diamond, overlapping rectangles, and cube). Scores range from zero to 11.
4. *Block Design* (Wechsler, 1981) is a test of visuoconstructional ability. The subject arranges colored blocks to match a stimulus design. Score reflects number correct with bonus points for rapid solution; range is from zero to 51.
5. *Word List Learning* (Morris et al., 1988) is a test of verbal memory. A 10-item word list is presented three times with free recall after each trial, followed by delayed free recall (range 0–10) and recognition (range 0–20) after a filled interval. Total learning is the number of words recalled across the three learning trials (range 0–30).
6. *Geriatric Depression Scale* (GDS; Yesavage et al., 1983) is a 30-item self-report measure of depression. Scores range from zero to 30. Performance of the patients and controls on these tasks is shown in Table 1.[2] In each case, higher scores indicate better function except on the GDS where higher scores indicate increasing depressive symptomatology.

This project was approved by the Indiana University institutional review board. All participants (or legal guardians) gave written informed consent and were paid a small sum of money.

Stimuli

For the Knowlton and Squire (1993) recognition and categorization tasks, we used one of the sets of nine-dot patterns originally used by these researchers. For the categorization task, the stimulus set consisted of a prototype, 20 low distortions of the prototype, 60 high distortions of the prototype, and 40 random patterns. The recognition stimulus set consisted of 10 unrelated or random patterns. For the third task, two new categories of dot patterns were generated using the method outlined by Posner et al. (1967). We chose two prototypes that were moderately similar to each other so as to ensure that the task would be a relatively demanding one. For each of these categories, we generated 7 low distortions (4 bits/dot), 7 medium distortions (6 bits/dot), and 16 high distortions (7.7 bits/dot) of the category prototype.

Procedure

All subjects participated in the three tasks described below. The order of the tasks was counterbalanced across participants. All tasks were conducted within a single session that lasted approximately 2 hr. The instructions for each task appeared on the screen and were read out loud by the experimenter. Participants were given the opportunity to briefly rest between each task.

Knowlton and Squire (1993) categorization task

In the training phase, participants viewed a single block of 40 different high distortions of a prototype. Each pattern

[1]In the early stage of data collection, 7 patients did not complete the experimental tasks. Five of these had particular trouble with the two-category task described below. We found that these patients had MMSE scores that were considerably lower than the rest of the group and so we targeted less demented patients for subject recruitment.

[2]Two amnestic subjects, Participant 4 and Participant 6, stand out in terms of their memory scores. As can be seen in Table 1, these individuals have memory scores that are significantly above the cut-offs. These patients both met the research diagnostic criteria for amnestic disorder at the time of testing but one of them was rediagnosed as normal in a follow-up visit. The phenomenon of “revert to normal” in these types of patients has been documented in large population studies (O’Connor et al., 1990; Unverzagt et al., 2001). Importantly, however, the general pattern of results does not differ when these two patients are omitted from the analyses, and none of our conclusions is changed.

Table 1. Characteristics of patient and control groups

Group	Participant	Age	Educ	MMSE	BN	CP	BD	WLL	WLD	GDS
Amnestic disorder	1	78	16	27	49	10	21	12*	3*	9
	2	69	11	28	51	9	20	14*	1*	4
	3	79	14	26	53	10	17	18	4*	4
	4	84	18	28	59	11	22	19	6	5
	5	71	8	28	37*	8	13	15	2*	8
	6	67	16	27	58	11	28	20	6	5
	7	82	12	25*	45	7*	16	12*	1*	6
	Mean	75.71	13.57	27.00	50.29	9.43	19.57	15.71	3.29	5.86
Dementia due to probable Alzheimer disease	8	74	11	22*	46	9	13	9*	2*	13
	9	75	8	24*	57	10	31	13*	3*	5
	10	71	12	20*	47	10	10	10*	1*	6
	11	81	12	29	50	10	19	13*	2*	4
	12	79	16	22*	57	11	19	12*	2*	1
	13	81	16	21*	28*	9	13	7*	1*	3
	14	79	12	25*	49	11	26	10*	0*	10
	15	78	12	23*	49	9	8	11*	2*	4
	16	74	8	21*	32*	11	22	12*	3*	8
	17	70	16	25*	55	9	25	16	3*	6
	Mean	76.20	12.30	23.20	47.00	9.90	18.60	11.30	1.90	6.00
Control	18	73	12	29	60	11	31	26	9	7
	19	77	16	29	59	11	25	18	8	5
	20	72	15	30	56	10	34	20	8	3
	21	70	18	30	57	11	27	24	9	2
	22	79	16	30	55	10	20	18	6	0
	23	76	13	29	58	10	23	18	5	7
	24	71	14	29	55	11	17	16	8	14
	25	70	13	30	57	11	36	25	9	7
	26	78	20	30	57	11	33	24	9	1
	27	73	19	27	55	11	23	21	7	4
	28	74	9	28	60	11	27	19	8	9
	29	71	12	30	58	11	32	16	8	1
	Mean	73.67	14.75	29.25	57.25	10.75	27.33	20.42	7.83	5.00

Note. Educ = Education. MMSE = Mini-Mental State Examination. Scores less than or equal to 25 are considered within an impaired range (Welsh et al., 1994). BN = Boston Naming. Scores less than or equal to 38 are considered impaired (Ivnik et al., 1996). CP = Constructional Praxis. Scores less than or equal to 7 are considered impaired (Welsh et al., 1994). BD = WAIS–R Block Design. Scores less than or equal to 10 are considered impaired (Ivnik et al., 2001). WLL = Word List Learning Sum Recall. Scores less than or equal to 14 are indicative of impairment (Morris et al., 1989). WLD = Word List Learning Delayed Recall. Scores less than or equal to 4 are considered impaired (Morris et al., 1989). GDS = Geriatric Depression Scale. Scores greater or equal to 15 are considered in depressed range (Yesavage et al., 1983). These cut-off scores are approximately 5th percentile of normal controls or 1.5 standard deviations below control means. GDS score is the mean for clinical patients with Mild Depression.
*Indicates values in the impaired range.

appeared on the screen for 5 s. Participants were told to simply attend to the patterns as they appeared on the screen. After the training phase was complete, the participants were told that the dot patterns all belonged to a single category in the same sense that if a series of dogs had been presented they would all have belonged to the category *dog*. The transfer test included 84 new patterns. In particular, the test included four instances of the category prototype, 20 low-level distortions, 20 high-level distortions, and 40 random patterns. Participants judged whether each pattern belonged to the same category as the patterns seen in training. The patterns were presented in a random order. Each pattern remained on the screen until the subject responded. No feedback was provided.

Knowlton and Squire (1993) recognition task

Participants in this task studied five random dot patterns eight times each. Again, in the training phase, participants were told to simply attend to the patterns as they appeared on the screen. Each pattern remained on the screen for 5 s. Within each of the eight blocks of study, the five items were presented in a random order. In the test that followed, participants saw the five old patterns and five new random patterns. Participants judged whether or not each pattern was exactly the same as one of the old patterns. Once again, the patterns remained on the screen until a response was collected, and the presentation order of the patterns was random. Again, no feedback was provided.

Two-category task

In this task, participants trained on two categories of dot patterns. For each category, the training set included two low distortions, two medium distortions, and six high distortions of the category prototype. For each level of distortion, the particular items that appeared as training items were randomly chosen from the full set of patterns for each participant. On a given trial, a pattern would appear on the screen and the prompt *Category A or B* would appear below the pattern. The participants' task was to classify the pattern. After each response they received feedback. If the response was correct, the word *correct* appeared. If the response was incorrect, the word *incorrect* appeared and the participants were told the correct category label for the item. There were 15 blocks of training. All training instances appeared once within each block. Order of presentation of the instances was randomized within each block.

After training was complete, participants were instructed that they would see some additional dot patterns, and that the task in this phase was to classify each pattern into one of the two categories. They were also told that unlike the first phase of the experiment, they would receive no feedback after their response. The transfer set included the original training items, four presentations of each prototype, five new low-level distortions, five new medium-level distortions, and 10 new high-level distortions of the prototype of each category. Order of presentation of the patterns in the transfer test was randomized for each participant.

RESULTS

As expected, the two patient groups performed similarly on the experimental tasks. In the Knowlton and Squire (1993) recognition task, the amnestic disorder patients ($M = .671$) performed similarly to the dementia due to probable Alzheimer's disease patients [$M = .640$; $t(15) = .350$, $p > .10$]. Following Knowlton and Squire (1993), in the categorization task a participant's response was scored as correct if the participant endorsed the prototype, low-level distortions, high-level distortions, or rejected the random patterns. In the Knowlton and Squire (1993) categorization task, there was no significant difference between the amnestic disorder patients ($M = .583$) and the dementia patients [$M = .552$; $t(15) = .820$, $p > .10$]. Finally, in the two-category dot-pattern task, these two groups performed similarly in the training phase [amnestic group: $M = .520$ *vs.* dementia group: $M = .556$; $t(15) = 1.30$, $p > .10$]; and the test phase [amnestic group: $M = .572$ *vs.* dementia group: $M = .549$; $t(15) = .539$, $p > .10$]. Therefore, in the remainder of the analyses, results will be combined across these groups.

The results of our replication of Knowlton and Squire's (1993) single-category task are shown in Figure 1A. The figure plots the probability with which the patient group and normal controls endorsed the different types of patterns as members of the category. Both groups show the classic typicality effect with the prototypes yielding the highest endorsement levels, followed in turn by the low distortions, high distortions, and random patterns. We conducted analyses of variances on these data using item type and group as factors. There was a main effect of item type on categorization responses [$F(3,81) = 14.55$, $p < .001$]. The control group and the patient group showed statistically equivalent categorization performance [$F(1,27) = .640$, $p > .10$]. In addition, there was no interaction of Group × Item Type [$F(3,81) = .811$, $p > .10$].

A composite summary of the results from the single-category task is shown in Figure 1B (leftmost bars), which plots the overall percent correct on the task. Critically, the patient group and the control group showed statistically equivalent accuracy on the categorization task [$t(27) = 0.817$, $p > .10$].

Figure 1B (rightmost bars) shows the overall percent correct for the patient group and normal controls in the recognition task. The normal controls performed significantly more accurately than did the patient group in this task [$t(27) = 2.284$, $p < .05$]. Furthermore, the interaction of Task (recognition *vs.* categorization) × Group (patient *vs.* control) was statistically significant [$F(1,27) = 5.178$, $p < .05$]. Therefore, these results replicate the dissociation between categorization and recognition observed by Knowlton and Squire (1993). Moreover, the results of the current study are numerically similar to those obtained by Knowlton and Squire (1993). For ease of comparison, Figures 1C and 1D display the data obtained by Knowlton and Squire (1993) for amnesics and controls in the same tasks.

Having verified that the patient group demonstrates the dissociation between categorization and recognition in Knowlton and Squire's (1993) tasks, we now turn to the critical question of interest, namely performance in the new two-category task. Figure 2 shows the probability correct for each of the training blocks in the two-category task for the normal controls and the patient group. Across the 15 blocks, the control group ($M = .593$) performed marginally better than the patient group [$M = .542$; $t(27) = 2.027$, $p = .053$]. The difference between the two groups is clear by the latter part of training. For example, in the final five blocks, the control group ($M = .628$) categorized the patterns with significantly higher accuracy than the patient group [$M = .554$; $t(27) = 2.059$, $p < .05$].

Figure 3 (top panel) shows the probabilities with which the various item types were endorsed by the patient group and the control group in the transfer test of the two-category task. The control group ($M = .643$) showed a general advantage over the patient group ($M = .559$) in categorizing the transfer patterns [$t(27) = 2.208$, $p < .05$]. An analysis of variance (Item Type × Group) restricted to the old patterns indicated that the control group was more accurate than the patient group at classifying training items in the transfer test [$F(1,27) = 5.021$, $p < .05$]. There was no main effect of item type or interaction of Item Type × Group for the old items. An analysis of variance (Item Type × Group) restricted to the new items indicated that the control

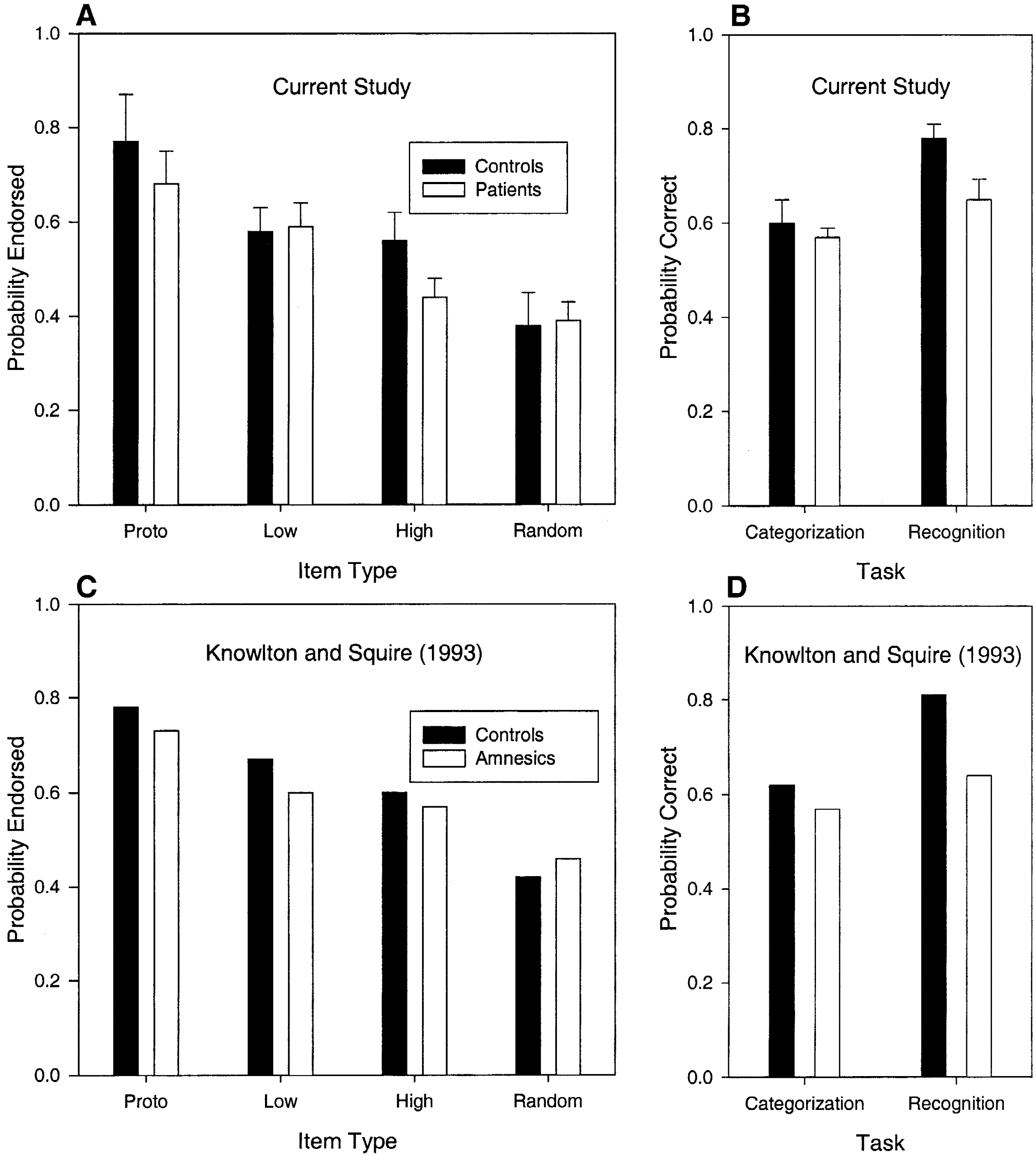

Fig. 1. Observed categorization and recognition performance in the current study and in Knowlton and Squire (1993). Panels A and C show the probability with which each item type was endorsed into the category by the control and patient groups in the current study and in Knowlton and Squire (1993), respectively. Panels B and D show the observed proportion of correct categorization and recognition decisions in the current study and in Knowlton and Squire (1993), respectively. Proto = prototype; Low = low-level distortions; High = high-level distortion, Random = random patterns. Error bars indicate the standard error of the mean.

group was also more accurate at classifying these patterns [$F(1,27) = 7.044$, $p < .05$]. The interaction of Item Type × Group was significant for the new items, indicating a typicality effect only for the control group [$F(3,83) = 4.436$, $p < .01$].

To compare performance on the new and old items, we conducted an ANOVA in which oldness and group were the two factors. In order for the distortion levels to be comparable across the old and new items, we restricted this analysis to the low-, medium-, and high-level distor-

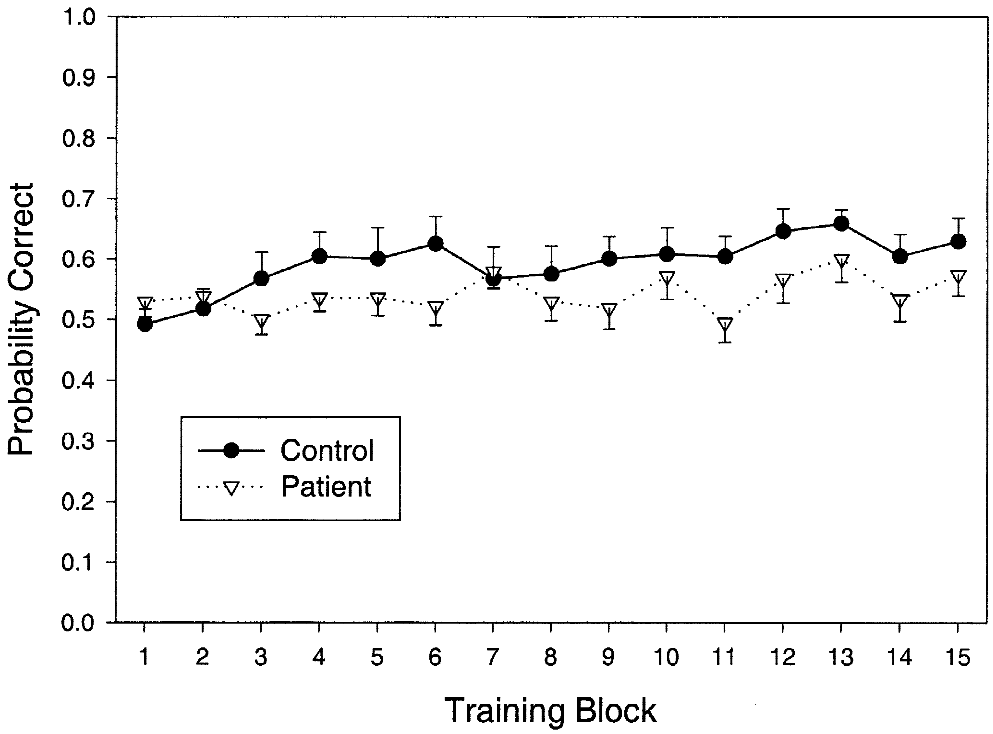

Fig. 2. Learning curves for the normal control and patient groups in the two-category experiment. Error bars indicate the standard error of the mean.

tions. Participants were generally more accurate at classifying the old items [$F(1,27) = 4.714$, $p < .05$]. However, there was no interaction of Oldness × Group [$F(1,27) = 0.619$, $p > .05$].

Finally, none of these results is changed if the data are rescored by assigning for each individual subject the category labels that would maximize performance. In other words, the pattern of results cannot be attributed to some individual subjects knowing the categories but reversing the labels.

The superior performance of the normal controls during the training portion of the two-category task, as well as on both old and new items during the transfer test, is consistent with the predictions from the single-system, exemplar-based account of categorization. Our interpretation is that in this more demanding categorization task, the patients' memory deficits were detrimental to their performance. Before turning to a fuller discussion of these results, however, we first consider the ability of the exemplar model to account quantitatively for the observed patterns of data in the two-category task.

EXEMPLAR MODEL ANALYSES

In this section we report quantitative fits of the exemplar model to the transfer data obtained in the two-category task. The model has already been shown in previous work to capture the pattern of results in Knowlton and Squire's (1993) single-category and recognition tasks (Nosofsky & Zaki, 1998; Nosofsky et al., 2001; Palmeri & Flanery, 1999).

According to the model, the probability that item i is classified in Category A is found by summing the similarity of the item to all Category A exemplars, and then dividing by the summed similarity of the item to all exemplars of both categories,

$$P(A|i) = \frac{\left[\sum s(i,a)\right]}{\left[\sum s(i,a)\right] + \left[\sum s(i,b)\right]} \tag{1}$$

where $s(i,a)$ denotes the similarity of item i to exemplar a. Following previous scaling and modeling work involving dot-pattern classification (Nosofsky et al., 2001; Smith & Minda, 2001), we computed similarity between exemplars as follows. First, the physical distance D between each pair of dot patterns is found by computing the average distance between each of the corresponding dots across the patterns.[3] This distance is then transformed by

[3] Correspondences between dots were established as follows. First, the dots of Prototype B were placed in correspondence with those of Prototype A by finding the one-to-one correspondence between the dots of the two patterns that minimized their computed distance. Thereafter, dot correspondences between the various distorted patterns were defined according to the dots' "movement histories." That is, two dots are placed in correspondence if they were perturbed from the same original dot in the prototype pattern.

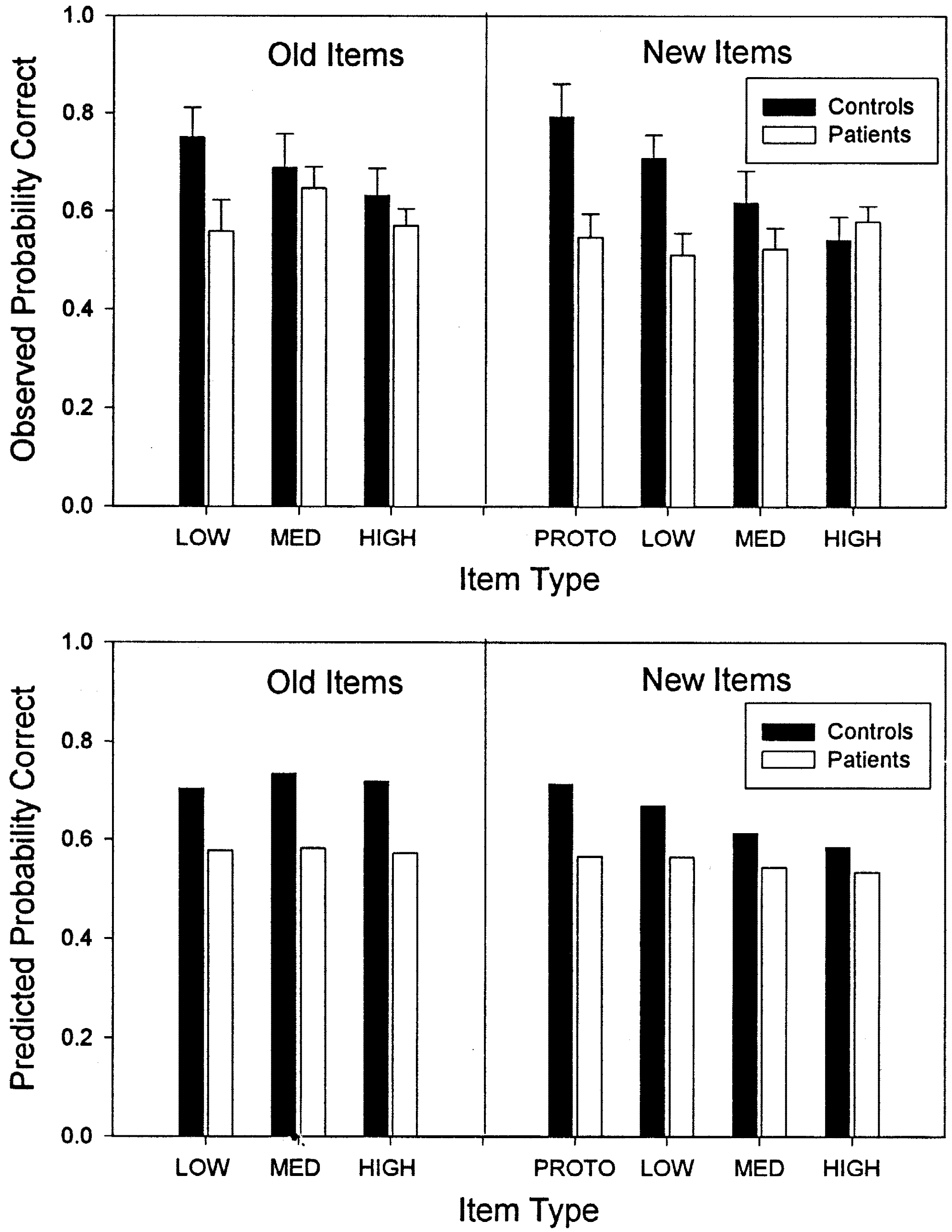

Fig. 3. Top panel: Observed correct categorization probabilities for the various types of transfer items in the two-category task. Bottom panel: Predicted probabilities from the exemplar model. Proto = prototype; Low = low-level distortions; Med = medium-level distortion; High = high-level distortions. Error bars indicate the standard error of the mean.

using the logarithmic transformation $d = \log(D + 1)$. Finally, the psychological similarity between the two patterns is assumed to be an exponential decay function of this transformed distance (Shepard, 1987), $s = \exp(-c \cdot d)$, where c is an overall memory sensitivity parameter. The computed similarities are then substituted into Equation 1 to yield the classification predictions from the model.

Fitting the model to the present data required the estimation of only two free parameters, the memory sensitivity parameter for the normal controls (c_n) and the memory sensitivity parameter for the patient groups (c_p). We conducted a computer search for the values of these free parameters that minimized the sum of squared deviations between the predicted and observed classification probabilities of the main item types.

The predicted probabilities from the exemplar model are displayed graphically in the bottom panel of Figure 3. The best-fitting parameters were $c_n = 1.083$ and $c_p = .429$. As expected, the estimated memory sensitivity of the patient group is substantially lower than that of the normal controls.

Although there is likely noise in the observed data due to the relatively small sample sizes, the exemplar model nevertheless captures many of the important trends (see Figure 3). First, it captures the result that the normal controls classified both the old and new transfer items with higher accuracy than did the patient group. Because the controls have greater memory sensitivity than do the patients, they are better able to discriminate between the exemplars of the contrasting categories. Second, the model captures the strong typicality gradient exhibited by the normal controls for the new transfer items. The reason is that the closer a transfer pattern is to the prototype, the more similar it tends to be to the training exemplars of its category. The model also predicts a much weaker new-item typicality gradient for the patients, because overall performance for this group is approaching a floor.

Third, the model predicts correctly that the old distortions are classified with higher overall accuracy than are the new distortions. It makes this prediction because each old distortion perfectly matches its own memory trace, which provides these patterns a performance boost. Finally, we were surprised by our earlier reported result that there was no statistically significant main effect of item type or Group $\times$ Item interaction for the old patterns. For technical reasons, however, it turns out that in the present design, the exemplar model predicts a nearly flat (slightly nonmonotonic) typicality gradient for the old items, so this result is in accord with the exemplar-model predictions as well.

DISCUSSION

Summary

In summary, previous work has reported dissociations in which amnesics perform substantially worse than do normal controls in explicit memory tasks such as recognition and recall, but perform at near-normal levels in tasks of categorization (Knowlton & Squire, 1993; Squire & Knowlton, 1995). Such dissociations support the idea that different representational systems underlie recognition and categorization. According to this interpretation, an explicit memory system, which is damaged in amnesics, mediates old–new recognition—whereas an implicit memory system, which is intact in amnesics, mediates categorization. However, researchers have also advanced single-system accounts of these dissociations. For example, Nosofsky and Zaki (1998) proposed that a single representational system based on stored exemplars underlies both categorization and recognition. According to this account, amnesics have a reduced ability to discriminate between distinct exemplars stored in memory. This reduced ability is represented in terms of a lowered level of a memory sensitivity parameter in the model. The key idea is that, in the tasks in which the dissociations have been demonstrated, the reduced level of memory sensitivity is extremely detrimental to old–new recognition performance, but is adequate to support near-normal levels of categorization.

The goal of the present research was to begin to distinguish between these alternative accounts. The key hypothesis from the single-system account was that if memory-impaired individuals were tested on a more demanding categorization task than has been used in previous work, then the dissociations between categorization and recognition should disappear. The past demonstrations of spared categorization performance in amnesics have primarily used single-category structures and did not include presentations of old training instances on the transfer tests. By contrast, in the present work, we used a two-category task in which observers needed to form representations of two separate categories in memory. Furthermore, to potentially increase the diagnosticity of the results, we included presentations of old training instances in the tests of category transfer.

Our main result was that in the previously employed single-category and recognition tasks, we replicated the dissociation between categorization and recognition observed by Knowlton and Squire (1993). However, on the more diagnostic two-category task, the memory-impaired participants performed significantly worse than did the matched normal controls. We interpret this result as support for the single-system hypothesis. We presented further support for the hypothesis by demonstrating that a single-system model based on the storage of individual exemplars gave a reasonably good quantitative account of the patterns of data in the two-category task, as long as it was assumed that the memory-impaired group had a lowered level of memory sensitivity compared to the group of matched normal controls.

Relations to Previous Work

Our experiment is not the first to test memory-impaired individuals on their ability to learn two-category problems. In this section we discuss the relation between the present research and some previous investigations. Knowlton et al. (1994) reported a study in which amnesics and matched normal controls engaged in tasks of probabilistic classification learning. In these tasks, four different cues were each probabilistically associated with one of two outcomes, and through trial-by-trial feedback the observers gradually learned these associations. The central result emphasized by Knowlton et al. (1994) was that during the first 50 trials of learning, the amnesics learned the associations at the same rate as the control subjects. However, at later stages of learning, the performance of the normal controls was significantly better than that of the amnesics. Knowlton et al. (1994) interpreted these results as evidence that an implicit memory system that is spared in amnesia is responsible for early learning in the probabilistic classification task, whereas declarative memory makes a contribution at later stages. However, Nosofsky and Zaki (1998) subsequently demonstrated that the single-system exemplar model described in this article predicts the learning results of Knowlton et al. (1994) without the need to posit separate memory systems. Intuitively, because at the very early learning stages very little exemplar-based information has been provided, the model predicts essentially no differences in performance between amnesics and normal controls during the first 50 trials. The results reported in the present experiment provide strong converging evidence consistent with the results from Knowlton et al. (1994) that impairments in declarative memory can result in significant deficits in performance at the later stages of two-category learning. Furthermore, our results show generality of these findings by demonstrating them in the dot-pattern classification task as well as in tests of classification transfer to new items.

Filoteo et al. (2001) trained amnesics and controls on a two-category perceptual classification task involving stimuli that consisted of horizontal and vertical lines of varying lengths connected in one corner. The category structure was defined by a complex quadratic rule. Filoteo et al. observed that accuracy for the last block of trials was comparable for controls and two amnesics. However, Filoteo et al. acknowledged that a simple strategy could have been adopted by participants to determine category membership in this task. The category was constructed so that if the two lines of the stimulus were approximately equal in length (i.e., looked like part of a square), then the item generally belonged to one category; whereas if the lines looked like part of a rectangle, then the item generally belonged to the other category. Filoteo et al. reported some analyses to rule out a simple model based on this explanation. However, Flanery et al. (2001) demonstrated that if the same category structure was mapped onto an alternative stimulus domain, then performance declined dramatically. The latter result is consistent with the idea that participants in Filoteo et al.'s study may indeed have relied extensively on a simple shape rule to complete the categorization task.

The previous multicategory experiment that is most closely related to the present one was conducted by Kolodny (1994), who also tested amnesics' classification performance in the dot-pattern paradigm. Indeed, Kolodny's (1994) results are often cited as further evidence in favor of a multiple-systems approach, because he observed that amnesics did not perform significantly worse than did normal controls in their dot-pattern classification. However, inspection of Kolodny's (1994) results indicates that his category learning data were numerically very similar to the data we reported here (see Kolodny, 1994, Figure 1). Throughout learning, the control subjects classified the dot patterns with roughly 10% higher accuracy than did the amnesic patients. These differences did not reach statistical significance, however, so Kolodny's (1994) data are catalogued as supporting a multiple-systems account. The sample sizes and number of training blocks were larger in the present study than in Kolodny's (1994) experiment, so the present design was a more powerful one for detecting a difference between groups.

Directions for Future Research

There are two important limitations of this research that need to be pursued in future work. The first concerns the idea that there may be multiple forms of categorization, only some of which are left unimpaired in amnesics. The second concerns the detailed nature of the memory impairment itself. We discuss each issue in turn.

The nature of the categorization task

Our memory-impaired patients performed worse than did the normal controls on the two-category task. This result is consistent with the prediction from the single-system model that as the memory demands of the categorization task are increased, performance deficits among amnesic subjects are more likely to be observed. However, a multiple-system theorist can argue that besides being more difficult, the two-category task differs in principle from the single-category tasks tested previously by Knowlton and Squire (1993) and Squire and Knowlton (1995). For example, the requirement that participants learn to associate a particular label with each category of objects may involve a form of declarative memory that is damaged in amnesics. Thus, an important avenue of future research would be to design alternative single-category paradigms that might reveal performance differences between normal and amnesic participants.

Nevertheless, as we reviewed in the previous section of this article, multiple-system theorists have sometimes advanced results from two-category learning paradigms in which amnesics did not perform significantly worse than did normal controls. If memory processes unique to two-category learning are indeed damaged in amnesics, then it is unclear why the amnesics did not display performance deficits in these previously reported studies. Our view is that, at present, the single-system model offers a more parsimonious account of the complete set of reported results.

The nature of the memory impairment

The memory-impaired patients that we tested in the present work seemed like a reasonable population to study, because these patients are known to have prominent deficits in various tasks of explicit memory. Conceivably, however, these patients may have suffered more widespread brain damage than the amnesics tested in the studies of Knowlton and Squire (1993) and Squire and Knowlton (1995), where the impairment was believed to be limited to the limbic-diencephalic brain system. In fact, it is known that the cognitive impairment of these patients, especially those diagnosed with dementia, is more widespread. For example, although many subjects in the two patient groups performed within the normal range on the various cognitive tests listed in Table 1, their performance on these tests was significantly different than the performance of the normal controls. Even the amnestic disorder participants, who by definition should have an isolated memory deficit, differ from the control group on a number of the cognitive tests. The more widespread impairment of these patients could be responsible for their performance deficits on the two-category task, so it is critical to pursue the present work by studying the behavior of other amnesic populations.

Nevertheless, it is important to remember that, on the tasks used by Knowlton and Squire (1993), these memory-impaired patients demonstrated the same dissociation between categorization and recognition as did the amnesics tested previously. However, on a more challenging categorization task, these patients show a definite categorization impairment. If the present results from the two-category task are to be discounted on grounds that these patients may not have the same neurological impairments as Knowlton and Squire's (1993) amnesics, then previous evidence involving dissociations between categorization and recognition on Knowlton and Squire's (1993) tasks should be discounted as well.

ACKNOWLEDGMENTS

This research was supported by a grant from the National Institute of Mental Health (PHS R01 MH48494) to R. Nosofsky, and a grant from the Indiana Alzheimer's Disease Center (P30 AG 10133) to B. Ghetti.

REFERENCES

American Psychiatric Association. (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington, DC: Author.

Filoteo, J.V., Maddox, W.T., & Davis, J.D. (2001). Quantitative modeling of category learning in amnesic patients. *Journal of the International Neuropsychological Society*, *7*, 1–19.

Flanery, M.A., Palmeri, T.J., & Shaper, B.L. (2001). Investigating dissociations between perceptual categorization and explicit memory. *Proceedings of the Twenty-Third Annual Conference of the Cognitive Science Society*, 291–296.

Folstein, M.F., Folstein, S.E., & McHugh, P.R. (1975). Mini-mental state: A practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research*, *12*, 189–198.

Goodglass, H. & Kaplan, E. (1983). *The assessment of aphasia and related disorders* (2nd ed.). Malvern, PA: Lea & Febiger.

Heit, E. (1993). Modeling the effects of expectations on recognition memory. *Psychological Science*, *4*, 244–252.

Hintzman, D.L. (1986). "Schema abstraction" in a multiple-trace memory model. *Psychological Review*, *93*, 411–428.

Hintzman, D.L. (1988). Judgments of frequency and recognition memory in a multiple-trace memory model. *Psychological Review*, *95*, 528–551.

Ivnik, R.J., Malec, J.F., Smith, G.E., Tangalos, E.G., Petersen, R.C., Kokmen, E., & Kurland, L.T. (1996). Neuropsychological tests' norms above age 55: COWAT, BNT, MAE Token, WRAT–R Reading, AMNART, STROOP, TMT, and JLO. *Clinical Neuropsychologist*, *10*, 262–278.

Ivnik, R.J., Malec, J.F.S., Smith, G.E., Petersen, R.C., Kokmen, E., & Kurland, L.T. (2001). Mayo's Older Americans Normative Studies: WAIS–R norms for ages 56 to 97. *Clinical Neuropsychologist*, *6* (Suppl.), 1–30.

Kinder, A. & Shanks, D.R. (2001). Amnesia and the declarative/nondeclarative distinction: A recurrent network model of classification, recognition, and repetition priming. *Journal of Cognitive Neuroscience*, *13*, 648–669.

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

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

Kolodny, J.A. (1994). Memory processes in classification learning: An investigation of amnesic performance in categorization of dot patterns and artistic styles. *Psychological Science*, *5*, 164–169.

McKhann, G., Drachman, D., Folstein, M., Katzman, R., Price, D., & Stadlan, E.M. (1984). Clinical diagnosis of Alzheimer's disease: Report of the NINCDS-ADRDA Work Group under the auspices of Department of Health and Human Services Task Force on Alzheimer's Disease. *Neurology*, *34*, 939–944.

Medin, D.L. & Schaffer, M.M. (1978). Context theory of classification learning. *Psychological Review*, *85*, 207–238.

Morris, J.C., Heyman, A., Mohs, R.C., Hughes, J.P., van Belle G., Fillenbaum G., Mellits E.D., & Clark C. (1989). The consortium to establish a registry for Alzheimer's disease (CERAD): I. Clinical and neuropsychological assessment of Alzheimer's disease. *Neurology*, *39*, 1159–1165.

Morris, J.C., Mohs, R.C., Rogers, H., Fillenbaum, G., & Heyman, A. (1988). Consortium to Establish a Registry for Alzheimer's Disease (CERAD) clinical and neuropsychological assessment of Alzheimer's disease. *Psychopharmacological Bulletin*, *24*, 641–652.

Nosofsky, R.M. (1986). Attention, similarity, and the identification-categorization relationship. *Journal of Experimental Psychology: General*, *115*, 39–57.

Nosofsky, R.M. & Zaki, S.R. (1998). Dissociations between categorization and recognition in amnesic and normal individuals: An exemplar-based interpretation. *Psychological Science*, *9*, 247–255.

Nosofsky, R.M., Zaki, S.R., & Palmeri, T.J. (2001). *Commentary on Smith and Minda's (2001) "Journey to the Center of the Category."* Manuscript submitted for publication.

O'Connor, D.W., Pollitt, P.A., Hyde, J.B., Fellows, J.L., Miller, N.D., & Roth M. (1990). A follow-up study of dementia diagnosed in the community using the Cambridge Mental Disorders of the Elderly Examination. *Acta Psychiatrica Scandinavica*, *81*, 78–82.

Palmeri, T.J. & Flanery, M.A. (1999). Learning about categories in the absence of training: Profound amnesia and the relationship between perceptual categorization and recognition memory. *Psychological Science*, *10*, 526–530.

Posner, M.I., Goldsmith, R., & Welton, K.E., Jr. (1967). Perceived distance and the classification of distorted patterns. *Journal of Experimental Psychology*, *73*, 28–38.

Posner, M.I. & Keele, S.W. (1968). On the genesis of abstract ideas. *Journal of Experimental Psychology*, *77*, 353–363.

Posner, M.I. & Keele, S.W. (1970). Retention of abstract ideas. *Journal of Experimental Psychology*, *83*, 304–308.

Reber, P.J. & Squire, L.R. (1999). Intact learning of artificial grammars and intact category learning by patients with Parkinson's disease. *Behavioral Neuroscience*, *113*, 235–242.

Reed, J.M., Squire, L.R., Patalano, A.L., Smith, E.E., & Jonides, J. (1999). Learning about categories that are defined by object-like stimuli despite impaired declarative memory. *Behavioral Neuroscience*, *113*, 411–419.

Shepard, R.N. (1987). Toward a universal law of generalization for psychological science. *Science*, *237*, 1317–1323.

Smith, J.D. & Minda, J.P. (2001). Journey to the center of the category: The dissociation in amnesia between categorization and recognition. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *27*, 984–1002.

Squire, L.R. & Knowlton, B.J. (1995). Learning about categories in the absence of memory. *Proceedings of the National Academy of Science*, *92*, 12470–12474.

Squire, L.R. & Zola, S.M. (1996). Structure and function of declarative and nondeclarative memory systems. *Proceedings of the National Academy of Science*, *93*, 13515–13522.

Unverzagt, F.W., Gao, S., Baiyewu, O., Ogunniyi, A.O., Gureje, O., Perkins, A., Emsley, C.L., Dickens, J., Evans, R., Musick, B., Hall, K.S., Hui, S.L., & Hendrie, H.C. (2001). Prevalence of cognitive impairment: Data from the Indianapolis Study of Health and Aging. *Neurology*, *57*, 1655–1662.

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

Welsh, K.A., Butters, N., Mohs, R.C., Beekly, D., Edland, S., Fillenbaum, G., & Heyman, A. (1994). The consortium to establish a registry for Alzheimer's Disease (CERAD): V. A normative study of the neuropsychological battery. *Neurology*, *44*, 609–614.

Yesavage, J.A., Brink, T.L., Rose, T.L., Lum, O., Huang, V., Adey, M., & Leirer, V.O. (1983). Development and validation of a geriatric depression screening scale: A preliminary report. *Journal of Psychiatric Research*, *17*, 37–49.

Zaki, S.R. & Nosofsky, R.M. (2001). A single-system interpretation of dissociations between recognition and categorization in a task involving object-like stimuli. *Cognitive, Affective, and Behavioral Neuroscience*, *1*, 344–359.