MEMORY SYSTEMS AND PERCEPTUAL CATEGORIZATION

Thomas J. Palmeri and Marci A. Flanery

I. Introduction

Do perceptual categorization and explicit memory depend on independent memory systems? Well-known exemplar models assume that judging whether some object belongs in a particular category—a categorization decision—and judging whether some object has been seen before—an explicit recognition memory decision—depend on the same underlying memory representations (e.g., Estes, 1994; Hintzman, 1986; Medin, 1986; Nosofsky, 1988, 1991; Nosofsky & Zaki, 1998). By contrast, many neuropsychological accounts assume that there are functionally independent memory systems subserving perceptual categorization and explicit memory (e.g., Squire & Zola, 1996). Evidence for multiple memory systems primarily comes from dissociations between categorization and explicit memory performance in studies of normals and amnesics. We review evidence from a variety of paradigms in which amnesics are reported to categorize at levels comparable to normals but are significantly impaired at explicit memory. Such dissociations appear to imply that separate systems may exist and seem to pose serious problems for theories that assume a single underlying memory system, such as exemplar models.

The evidence is clear that amnesics have impaired explicit memory. The focus of this paper is on whether data from studies testing amnesics provide similarly clear evidence for completely intact memories for newly learned perceptual categories.

THE PSYCHOLOGY OF LEARNING
AND MOTIVATION, VOL. 41

0079-7421/02 $35.00

We will also discuss whether models assuming a single memory system can account for observed dissociations between categorization and explicit memory. We will review some of the behavioral evidence for multiple independent memory systems and in each case will describe some recent work that challenges the conclusions of these various studies.

II. Dot Pattern Classification Studies

A classic methodology for studying categorization and recognition has been the Posner and Keele (1968, 1970) dot pattern paradigm (e.g., Homa, 1984; Knowlton & Squire, 1993; Nosofsky & Zaki, 1998; Palmeri & Nosofsky, 2001; Shin & Nosofsky, 1992). To create a dot pattern, a small number of dots are randomly scattered on a grid. To create a category, one pattern is randomly generated and designated the category prototype. Category members are generated by randomly distorting the prototype by moving each dot in the prototype in a random direction by an amount proportional to the degree of distortion desired (Posner, Goldsmith, & Welton, 1967). Figure 1 displays a category prototype, a low-level distortion, a high-level distortion, and a randomly generated nonmember.

A. A Dissociation between Categorization and Recognition

Knowlton and Squire (1993) adapted a variant of this paradigm to test amnesics and normals on categorization and recognition. In the categorization task, subjects

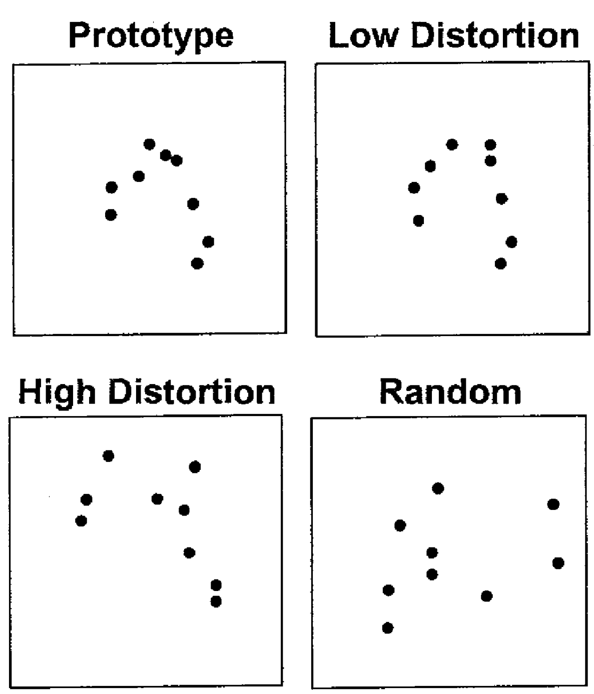

Fig. 1. Examples of a prototype, low distortion, high distortion, and random dot pattern used in dot pattern classification studies. (Stimuli from Knowlton & Squire, 1993.)

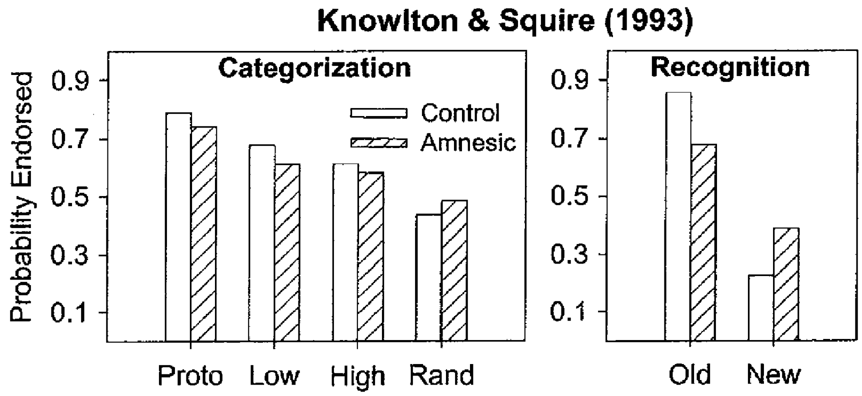

Fig. 2. Categorization and recognition memory data for controls and amnesics from Knowlton and Squire (1993). The left panel displays the observed probability of endorsing prototypes (Proto), low distortions (Low), high distortions (High), and random patterns (Rand) as members of the studied category. The right panel displays the observed probability of endorsing old and new items as old stimuli. (From Palmeri, T. J., and 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**, pp. 526–530.)

were initially exposed to 40 high-level distortions of the category prototype. An implicit learning task was used in which subjects were simply asked to point to the center dot of each pattern. After a 5-min delay, subjects were told that the patterns all belonged to the same category and were asked to judge new patterns as members or nonmembers of that category. Category members consisted of 4 repetitions of the category prototype, 20 low-level distortions, and 20 high-level distortions. Nonmembers were 40 randomly generated patterns. Dot patterns were presented one at a time, in random order, and subjects were asked to judge each pattern as a member or nonmember of the previously viewed category without corrective feedback. The left panel of Fig. 2 displays the probability of endorsing the prototypes, low distortions, high distortions, and random patterns as category members for amnesics and age-matched normal controls. Knowlton and Squire observed that amnesics were not significantly worse at categorization than normal controls.

In the recognition memory task, subjects were exposed to five randomly generated patterns eight times each (thus equating for the number of exposure trials used in the categorization task). As in the categorization task, subjects were asked to point to the center dot of each pattern without being told that they would later be tested on their memory for the dot patterns. After a 5-min delay, subjects were asked to discriminate between the five old patterns and five new patterns. Again, no corrective feedback was provided. As shown in Fig. 2, a behavioral dissociation was observed in that amnesics were significantly impaired at discriminating old from new patterns in the recognition memory task, but were not significantly impaired at discriminating members from nonmembers in the categorization task.

This pattern of results has been used as evidence for two independent memory systems: an explicit hippocampal-dependent declarative memory system subserving recognition memory, which is impaired in amnesia, and an independent implicit categorization system, which is spared in amnesia. Knowlton and Squire (1993, p. 1748) concluded that "single-factor models in which classification judgments derive from, or in any way depend on, long-term declarative memory do not account for the finding that amnesic patients perform well on the classification tasks."

B. A Single-System Exemplar Account

Although these results seemed to demonstrate the existence of independent systems for categorization and recognition, Nosofsky and Zaki (1998) reported theoretical analyses that showed a single-system exemplar model capable of accounting for this empirical dissociation in a fairly straightforward manner. By simply assuming that amnesics had poorly discriminated memory traces (low memory sensitivity) compared to normals, which was instantiated by variation in a single parameter of the model, the exemplar model was able to account for the observed dissociation between recognition and categorization (see also Nosofsky, 1988). As shown in Fig. 3, simulations with a high value of memory sensitivity (high c) generated predictions comparable to observed behavior by normal controls and simulations with a low value of memory sensitivity (low c) generated predictions comparable to observed behavior by amnesics.

C. An Extreme Dissociation

One important factor that allows the exemplar model to successfully account for the Knowlton and Squire (1993) results is that amnesics had poor but above-chance

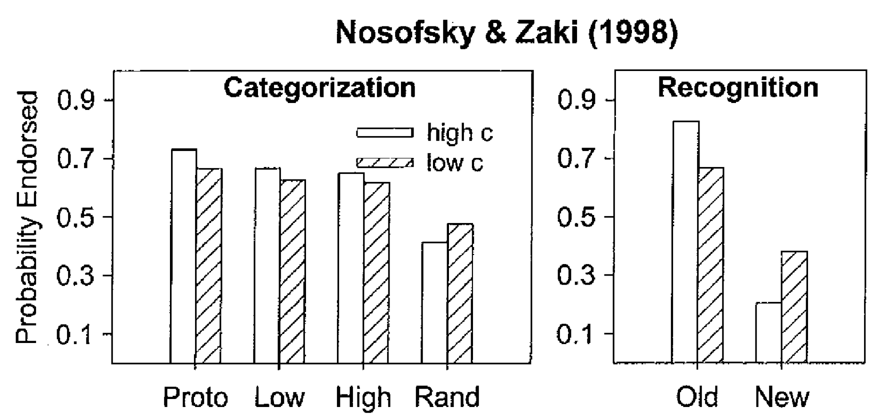

Fig. 3. Categorization and recognition memory predictions for an exemplar model with varying levels of memory sensitivity (high c versus low c) from Nosofsky and Zaki (1998). The left panel displays the predicted probability of endorsing prototypes (Proto), low distortions (Low), high distortions (High), and random patterns (Rand) as members of the studied category. The right panel displays the predicted probability of endorsing old and new items as old stimuli.

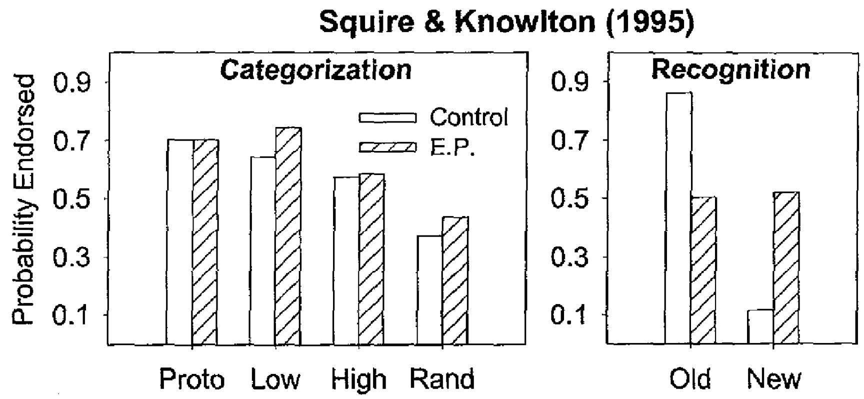

Fig. 4. Categorization and recognition memory data for controls and the profound amnesic E.P. from Squire and Knowlton (1995). The left panel displays the observed probability of endorsing prototypes (Proto), low distortions (Low), high distortions (High), and random patterns (Rand) as members of the studied category. The right panel displays the observed probability of endorsing old and new items as old stimuli. (From Palmeri, T. J., and 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**, pp. 526–530.)

recognition memory; as shown in Fig. 3, the model predicted very small deficits in categorization but far larger deficits in recognition memory. Other evidence reported by Squire and Knowlton (1995) may be more challenging to single-system models. They tested E.P., a profoundly amnesic individual, on tasks very similar to those used by Knowlton and Squire (1993). As shown in Fig. 4, as with other amnesic individuals, E.P. was able to categorize as well as normals. However, E.P. was completely unable to recognize old versus new patterns better than chance. In summarizing these results, Squire and Zola (1996) concluded that

> these results suggest that category knowledge can develop independently of and in the absence of normal declarative memory . . . the information supporting classification learning must be distinct from declarative knowledge about the specific items presented for training. Models in which classification judgments derive from, or in any way depend on, long-term declarative memory do not account for the finding that amnesic patients can acquire category knowledge as well as normal subjects. (pp. 13,517–13,518)

Indeed, as illustrated later, it may prove quite challenging for a single-system exemplar model to account for this extreme dissociation without some augmentation (see Nosofsky & Zaki, 1998); in order for the exemplar model to predict chance recognition memory performance, it must predict chance categorization performance as well.

D. Reevaluating the Experimental Paradigm

Squire and Knowlton's (1995) findings may appear devastating to the single-system models. However, we have argued that the experimental procedures used to test

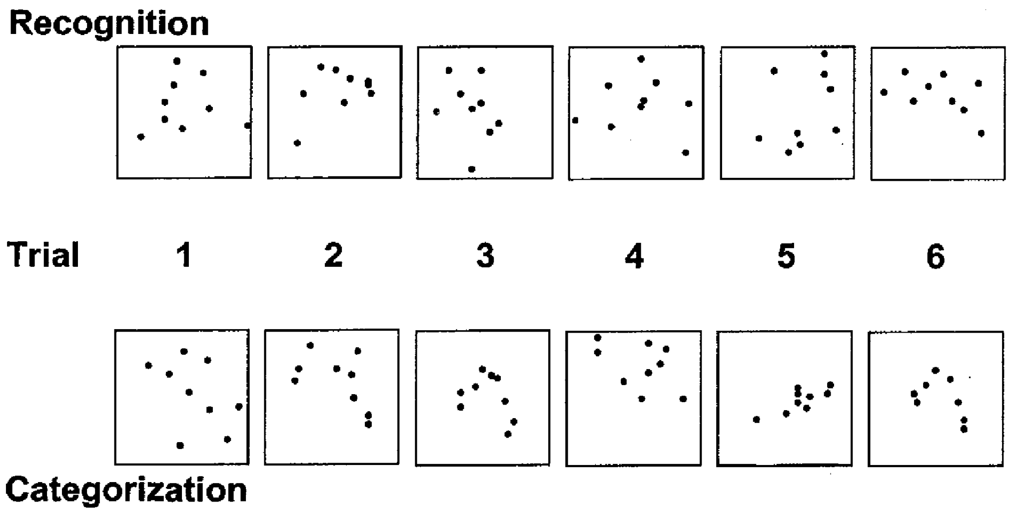

Fig. 5. Example sequences of recognition trials (top row) and categorization trials (bottom row) from Knowlton and Squire (1993). For recognition, Trials 1, 3, and 5 show old patterns, and Trials 2, 4, and 6 show new patterns. For categorization, Trials 1, 4, and 5 show nonmembers, and Trials 2, 3, and 6 show category members; Trial 2 shows a high-level distortion of the prototype, Trial 3 shows the prototype, and Trial 6 shows a low-level distortion of the prototype. (From Palmeri, T. J., and 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**, pp. 526–530.)

E.P. and other amnesics may be fundamentally flawed in that prior exposure to training stimuli is unnecessary to accurately perform the categorization task (Palmeri & Flanery, 1999). To illustrate, the top row of Fig. 5 displays a sequence of recognition memory test trials from Knowlton and Squire (1993). Not surprisingly, it is impossible to judge which of these patterns are old or new without ever having seen the training patterns. The bottom row of Fig. 5 displays a sequence of categorization trials. Recall that category members are the prototype, low distortions of the prototype, and high distortions of the prototype, and that nonmembers are a set of entirely random patterns. As may be apparent from the figure, without any prior exposure to the category, it is possible to discover that a set of very similar patterns all belong to the same category and that a set of very dissimilar patterns are all nonmembers of that category. In fact, such judgments should be possible in the absence of much if any long-term memory for the patterns. Thus, a profound amnesic, such as E.P., who has otherwise normal cognitive functioning, apart from his profound declarative memory deficit, may be able to judge category membership without much if any memory for the previously studied patterns.

1. *Learning about Categories in the Absence of Training*

Palmeri and Flanery (1999) investigated whether prior exposure was even necessary to categorize the test items. Again, one explanation for above-chance categorization by amnesics is that it may be possible to group test items that look similar (the prototype and its distortions) into the member category and group test items

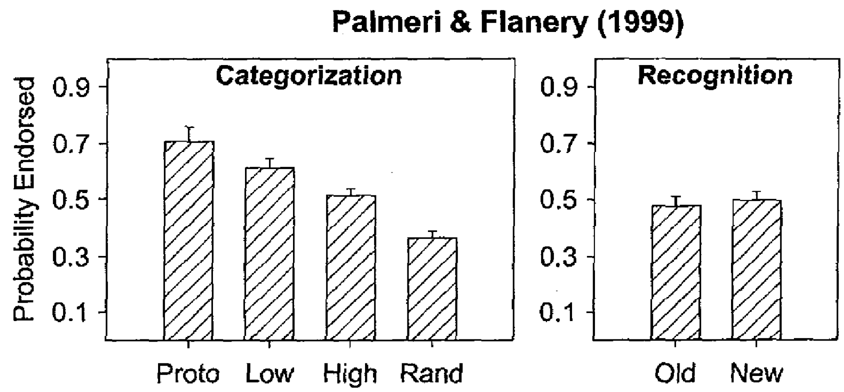

Fig. 6. Categorization and recognition data for simulated amnesics from Palmeri and Flanery (1999). The left panel displays the observed probability of endorsing prototypes (Proto), low distortions (Low), high distortions (High), and random patterns (Rand) as category members. The right panel displays the observed probability of endorsing old and new items as old stimuli. (From Palmeri, T. J., and 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**, pp. 526–530.)

that do not look similar (random patterns) into the nonmember category. By contrast, it is impossible to tell apart old from new patterns without memory. Palmeri and Flanery tested this possibility by producing a state of profound amnesia in normal subjects. As a ruse, subjects were told that patterns had been subliminally presented during an initial word identification task. In fact, no dot patterns were ever really presented. Subjects then completed the same categorization and recognition tests used by Knowlton and Squire (1993). As shown in Fig. 6, like E.P., our simulated profound amnesics showed chance recognition, as expected. Yet, our subjects showed above-chance categorization. Indeed, our simulated profound amnesics were 60.4% correct at categorizing dot patterns, performance that was in close correspondence to that observed by amnesics (59.9%, Knowlton & Squire, 1993), by E.P. (61.1%, Squire & Knowlton, 1995), and by college students after a 1-week delay (57%, Nosofsky & Zaki, 1998). Apparently, our subjects were able to categorize members versus nonmembers by picking up on the category structure clearly embedded within the categorization test. They had no prior memories for training items to rely on. Indeed, when debriefed at the end of the experiment, some of our subjects insisted that they must have seen dot patterns during the "subliminal exposure" phase of the categorization task since they were able to categorize the test items with such confidence.

2. *Experiment 1: How Much Information Can Be Acquired in the Absence of Training?*

So, even without memory for the category members, it may be possible to correctly categorize members versus nonmembers in the particular type of dot pattern paradigm used by Knowlton and Squire (1993; Squire & Knowlton, 1995).

Our first study demonstrated that subjects achieved around 60% accuracy at judging members versus nonmembers without the benefit of any prior exposure to the category. Although comparable to the performance of amnesics and normal individuals reported in other studies, one might argue that the amount of information that can be acquired in the absence of prior exposure to category members might be relatively meager, permitting classification performance that is barely better than chance. As a way of maximally assessing how much information could possibly be extracted from the categorization test sequence, a particularly well-motivated and informed subject (the second author) participated in 10 categorization test sessions, with a new computer-generated set of stimuli used within each session. These categorization tests had the exact same abstract structure as those used by Knowlton and Squire (1993). However, in our experiment, the subject did not receive any prior exposure to category members. Although she was aware of how the category members and nonmembers were defined abstractly, she had absolutely no prior knowledge of the particular prototypes and distortions that were to be used within a given test session—that is, she needed to discover which patterns were members or nonmembers without the benefit of any prior exposure and without the benefit of any corrective feedback. It is important to emphasize that even with a complete understanding of the procedures for how old patterns and new patterns were generated in a recognition test, it would be absolutely impossible to recognize old from new patterns better than chance without having seen old patterns before. From the perspective of a potential subject without any prior exposure, even one who is particularly motivated and well informed, the patterns used during a recognition memory test are assigned as "old" or "new" patterns in a completely arbitrary manner.

As shown in Fig. 7, without any prior exposure to category members, this subject was able to correctly categorize the prototypes perfectly, the low distortions nearly perfectly, and the high distortions and random patterns extremely well, achieving an overall accuracy of 81.3% correct. As we expected, there is a tremendous

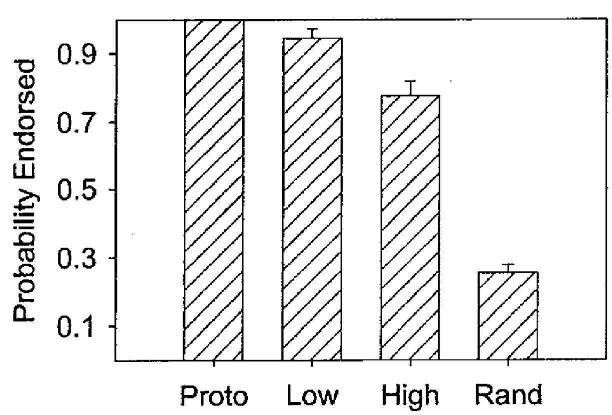

Fig. 7. Categorization data for a single motivated subject who completed the categorization test without prior exposure to category members (Experiment 1). The figure displays the observed probability of endorsing prototypes (Proto), low distortions (Low), high distortions (High), and random patterns (Rand) as members of the category.

amount of information in the sequence of categorization trials for a particularly well-motivated subject to use to accurately categorize the test stimuli from the Knowlton and Squire (1993) paradigm without any prior exposure to category members. This result is troublesome because in testing the relative independence or nonindependence of categorization and recognition, it is critical that the tasks be equated for how much they actually depend on memory for training patterns presented earlier in the experiment. The proper categorization test would mirror the recognition memory test in that accurate performance would be impossible without prior exposure to category members.

3. *Experiment 2: What about Prior Training?*

The above studies clearly demonstrate accurate categorization in the absence of training. But what happens when our subjects are actually given prior exposure to high-level distortions of a category prototype? A straightforward hypothesis is that subjects who receive prior exposure to category members should be able to classify significantly more accurately than subjects who receive no prior exposure. However, such a finding could seriously undermine our claim that simulated amnesics can be used to understand the classification abilities of true amnesics. Our claim is that amnesics may base their categorization responses on information acquired during the testing session. Yet, amnesics do not classify significantly worse than normal individuals, who presumably can use their memory for the category members that they were shown just a few minutes earlier.

We directly compared the performance of subjects who were actually exposed to the study items (Exposure) to that of simulated amnesics (None). That is, half of the subjects were given subliminal exposure, as in Palmeri and Flanery (1999), and were then tested on categorization or recognition memory; the other half were given actual exposure, as in Knowlton and Squire (1993), and were then tested on categorization or recognition memory. Figure 8 displays response accuracy in the categorization and recognition conditions in this experiment and displays average results from the experiments with amnesics and normal controls for comparison. As expected, the exposure group could recognize items well above chance but the no-exposure group could only guess. Replicating Palmeri and Flanery (1999), subjects in a no-exposure group could categorize well above chance. Interestingly, subjects receiving no exposure did not categorize significantly worse than subjects who were actually exposed to category items. Apparently, in this particular paradigm, prior exposure to a category does not provide much, if any, benefit for categorizing items later. Although perhaps surprising, we should remind readers that the study items in this particular paradigm were all high-level distortions of the category prototype. These items do not look very similar to the category prototype nor do they look similar to one another. Indeed, in most experiments using this paradigm, high-level distortions are typically rated for category membership at levels around 50%.

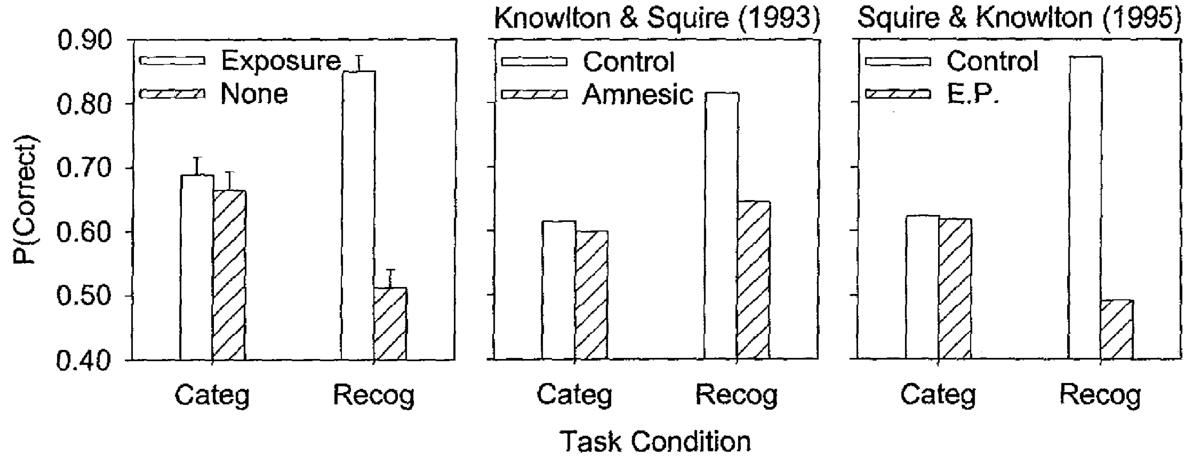

Fig. 8. Percentage correct categorization and recognition from Experiment 2, Knowlton and Squire (1993), and Squire and Knowlton (1995). The Palmeri and Flanery results display categorization and recognition accuracy as a function of prior exposure to category members or old items, respectively. The Knowlton and Squire results display categorization and recognition memory accuracy for controls and amnesics. Correct categorization decisions were defined as judging the prototype, low distortions, and high distortions as members and judging random patterns as nonmembers. Correct recognition decisions were defined as judging old patterns as old and new patterns as new.

4. *Experiment 3: Examining Different Kinds of Prior Category Exposure*

One potential criticism of the studies described above is that the ruse used to induce amnesia may actually place subjects in a very different cognitive set from that of subjects who were actually exposed to category members. In other words, our "profound amnesics" may realize that they never saw any patterns at all and may think that the task is to discover the hidden category structure, something they appear to do quite ably. So, one goal of the following experiment was to use a different paradigm for demonstrating that subjects may categorize based on information they acquire during the categorization test. As described later, in this experiment, we surreptitiously switched the test stimuli for some subjects to that of an unstudied category in order to see if they would categorize test stimuli based on what they had studied earlier or if they would instead categorize test stimuli based on the information presented within the categorization test.

In addition, we clearly do not want to draw the erroneous conclusion that people always ignore information about a previously studied category in favor of information presented during a categorization test. Therefore, a second goal was to show that when initial exposure provides clear evidence for a category structure, subjects will use that information to make category decisions irrespective of the makeup of the categorization test. To demonstrate this, we adapted additional aspects of the paradigm used by Squire and Knowlton (1995). In one condition, subjects were initially exposed to 40 high distortions of the prototype (40H), exactly as was done in all of the earlier studies. In another condition, subjects were instead exposed to 40 repetitions of the category prototype (40P). We reasoned that subjects in the 40P condition should have acquired clear knowledge of the category structure and should protest any surreptitious changes during a categorization test. By contrast, subjects in the 40H condition should have acquired little knowledge of the category structure and should go along with our surreptitious changes.

First, in order to verify that different exposure conditions had a significant effect on performance, we tested subjects in the same way as we did in our earlier studies after a 1-week delay. Overall, 40P subjects achieved 77.2% accuracy and 40H subjects achieved 64.0% accuracy. As expected, categorization accuracy was significantly influenced by the type of information presented during initial category exposure, as was reported by Squire and Knowlton (1995) for normal subjects. Overall performance of our 40H subjects was quite comparable to what we and others have observed in this paradigm; performance of the 40P subjects was significantly better than what we have observed before (but was comparable to the performance by our single motivated subject in Experiment 1 described in Section II.D.2). So, information presented during initial exposure can have a significant effect on categorization performance, as we predicted.

As a way of simulating amnesia, we tested these subjects after an additional delay of several weeks (see Nosofsky & Zaki, 1998). But now we tested just half of the subjects on items generated from the prototype used to generate items they had seen before (Same condition) and tested the other half of the subjects on items generated from a novel prototype (Different condition). Thus, each subject was assigned to one of four conditions: 40P-Same, 40P-Different, 40H-Same, and 40H-Different. Because all subjects were given different randomly generated stimuli, we can characterize subjects in the Different condition as mistakenly receiving a categorization test that was intended for another individual.

As illustrated in Fig. 9, we found that subjects in the 40P-Same condition performed quite well, correctly categorizing over 70% of the items. However, subjects in the 40P-Different condition were completely at chance categorizing the test items. We suspect that these subjects tried to use the category information they had clearly acquired earlier and could not apply that knowledge when given a test comprised of entirely novel items. By contrast, for subjects in the 40H conditions, there was no significant difference in performance between subjects who were tested on the same category structure they were initially exposed to and subjects who were tested on a completely novel category structure. Consistent with our previous results, these subjects appear to be making categorization decisions based on information acquired during the categorization test, not on what they may have acquired during earlier phases of the experiment.

5. *Summary*

The dissociation between categorization and recognition reported by Knowlton and Squire (1993; Squire & Knowlton, 1995) initially appeared to present strong

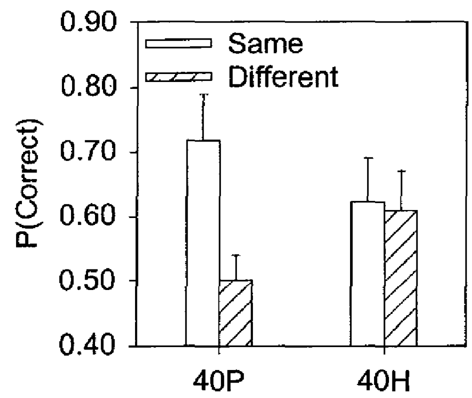

Fig. 9. Percentage correct categorization as a function of studied category (40P versus 40H) and as a function of categorization test (Same versus Different) from Experiment 3. 40P subjects studied 40 repetitions of the prototype. 40H subjects studied 40 high distortions. Same-condition subjects were tested on stimuli generated from the prototype used to generate their studied category items. Different-condition subjects were tested on stimuli generated from a novel prototype. Correct categorization decisions were defined as judging the prototype, low distortions, and high distortions presented during the test as members and judging random patterns presented during the test as nonmembers.

evidence supporting multiple memory systems theory. Our experiments reported how the observed dissociation between categorization and recognition using distorted dot patterns may be explained as a result of the particular methodologies used to test these individuals. We showed that very good categorization performance can be achieved in the absence of any prior exposure to the category members, and that this performance is comparable to that of subjects who had been provided prior exposure. We also showed that very good categorization performance can be achieved when people are tested on items that are different from what they had actually studied. But this seems to only occur when subjects have been initially exposed to a very diffuse category structure consisting of high distortions that are not very similar to one another, which was also true of the experiments used by Knowlton and Squire (1993; Squire & Knowlton, 1995). When subjects have been exposed to a clear category structure through repetition of a single prototype, they attempt to categorize items based on that acquired category knowledge, not on information presented during the categorization test.

III. Theoretically Modeling Dot Pattern Classification

Our focus will now shift to examining how formal models of categorization have attempted to account for the dissociation between categorization and recognition observed by Knowlton and Squire (1993). For this discussion, we will just make the assumption that subjects acquire information about a category during an initial study session and then utilize that acquired category knowledge during the categorization test. We will forgo considerations of how category information may be acquired during the categorization test itself until later in this section.

A. An Exemplar-Based Interpretation

According to exemplar models, categories are represented in terms of stored category exemplars (Hintzman, 1986; Medin & Schaffer, 1978; Nosofsky, 1986). For example, in order to theoretically model behavior in the Knowlton and Squire (1993) paradigm, the studied category is assumed to be represented in terms of the 40 high distortions (the exemplars) of the category prototype. The total evidence that an item presented during the categorization test is a member of that studied category is given by the summed similarity to the stored exemplars of the category. As described by Nosofsky and Zaki (1998), when just a single category is learned, the probability of classifying item i as a member of the studied category M is given by

$$P(M \mid i) = \frac{\sum_{m \in M} s_{im}}{\sum_{m \in M} s_{im} + k_C} \tag{1}$$

where s_{im} is the similarity between item i and stored exemplar m, and k_{C} is the response criterion for categorization. According to this equation, if the summed similarity is greater than k_{C}, then the probability of classifying item i as a member of category M will be greater than .50.

In contrast to multiple memory systems theories, a fundamental assumption of exemplar models is that categorization and recognition depend on the same underlying exemplar memories. So, analogously, the probability of judging an item as being an old item in the recognition memory task is found by summing the similarity to the studied exemplars and comparing this summed similarity to a recognition criterion. Specifically, the probability of judging item i to be an old item is given by

$$P(\mathrm{old} \mid i) = \frac{\sum_{m \in old} s_{im}}{\sum_{m \in old} s_{im} + k_R} \tag{2}$$

where k_{R} is the response criterion for recognition.

In the full version of the generalized context model (Nosofsky, 1984, 1986), the similarity between item i and stored exemplar m is given by

$$s_{im} = \exp(-c \cdot d_{im}) \tag{3}$$

where d_{im} is the distance between item i and exemplar m in similarity space, and c is the sensitivity parameter. In previous applications of the exemplar model to dot pattern experiments, multidimensional scaling has been used to derive the distances, d_{im}, between patterns in psychological space (e.g., Palmeri & Nosofsky, 2001; Shin & Nosofsky, 1992). Unfortunately, with 40 training items and 84 test items, it would require thousands of pairwise similarity ratings to derive the underlying similarity space. So, for obvious practical reasons, Nosofsky and Zaki (1998) just obtained a subset of similarity ratings between different types of patterns from each individual subject. Specifically, they obtained average similarity ratings between old high distortions and the prototype, low distortions, new high distortions, and random patterns; to model the recognition memory results, they also obtained average similarity ratings between pairs of random patterns (recall that all old and new items in the recognition memory experiment were random patterns). As a simple approximation, they assumed that the true psychological similarity between different types of patterns was given by a power transform of their rated similarity

$$s_{im} = [\mathrm{rating}(i, m)]^p \tag{4}$$

where rating(i, m) is the average rated similarity between an item of type i and an item of type m. Moreover, using the relation that $\exp(-\mathrm{c} \cdot \mathrm{d}) = [\exp(-\mathrm{d})]^{\mathrm{c}}$, Nosofsky and Zaki (1998) noted that increases in the value of the sensitivity parameter, c, could be modeled by increases in the value of exponent p in Eq. (4), a point that will be critical in the ensuing discussion.

Combining the above equations, the probability of classifying item i as a member of category M is given by

$$P(M \mid i) = \frac{40 \times [\mathrm{rating}(i, h)]^p}{40 \times [\mathrm{rating}(i, h)]^p + k_C} \tag{5}$$

where rating(i, h) is the average similarity rating between an item of type i and an old high distortion. The probability of judging old item i as an old item is given by

$$P_{old}(\mathrm{old} \mid i) = \frac{\delta_{ii}^p + 4 \times [\mathrm{rating}(r, r)]^p}{\delta_{ii}^p + 4 \times [\mathrm{rating}(r, r)]^p + k_R} \tag{6}$$

where δ_{ii} is the self-similarity between old item i and its own stored representation (a free parameter) and rating(r, r) is the average similarity rating between two random patterns. And the probability of judging new item i as an old item is given by

$$P_{new}(\mathrm{old} \mid i) = \frac{5 \times [\mathrm{rating}(r, r)]^p}{5 \times [\mathrm{rating}(r, r)]^p + k_R} \tag{7}$$

The key insight by Nosofsky and Zaki (1998) was to consider the possibility that the observed behavioral dissociation, in which amnesics could categorize quite well but were significantly impaired at recognition memory, could reflect a single parameter difference between amnesics and normals. Impaired memory in amnesia could be simulated by a difficulty in discriminating between exemplars in memory. Conceptually, in the similarity calculations given by Eq. (3), a relatively high value of c (high memory sensitivity) causes memories to be easily discriminated from one another, but a low value of c (low memory sensitivity) causes memories to be much less discriminable from one another. Specifically, Nosofsky and Zaki (1998) found that the power parameter p (which reflects values of memory sensitivity) in Eq. (4) was larger for simulated normals than for simulated amnesics, indicating a lower level of memory sensitivity in amnesia. As shown previously in Fig. 3, Nosofsky and Zaki (1998) demonstrated that this single parameter difference between amnesics and normals allowed a single-system exemplar model to account for the dissociation between categorization and recognition reported by Knowlton and Squire (1993).

B. A Prototype-Based Interpretation

By contrast, Knowlton and Squire (1993) interpreted the dissociation between categorization and recognition in terms of independent memory systems. Recognition judgments are determined by a declarative memory system based on the storage of individual exemplars, which is clearly damaged in amnesia. Categorization judgments are determined by an implicit memory system based on the formation of abstract prototypes, which is apparently spared in amnesia.

In support of the multiple memory systems view, Smith and Minda (2001) recently provided an extensive critique of the Nosofsky and Zaki (1998) article. The emphasis of their critique was that Nosofsky and Zaki collected direct similarity ratings between dot patterns after they had completed both category training and transfer. Smith and Minda instead proposed using an "objective" measure of similarity between two patterns based on physical distances between the individual dots in the two patterns (Posner et al., 1967). Specifically, assume that D_{im} is the average physical Euclidean distance between the dots in presented item i and the dots in stored item m. Smith and Minda then assumed that this average distance was log transformed according to $d_{\text{im}} = \log(D_{\text{im}} + 1)$. The similarity between item i and stored item m is then given by $s_{\text{im}} = \exp(-c \cdot d_{\text{im}})$, which is just Eq. (3). In fitting the exemplar model, Eq. (1) was used. In fitting a multiplicative prototype model, an analogous equation was used

$$P(M \mid i) = \frac{s_{iP}}{s_{iP} + k_C} \tag{8}$$

where s_{iP} is the similarity between item i and the category prototype P. In fitting the prototype model to the observed data, Smith and Minda assumed that the stored category prototype was the population-based prototype. The population-based prototype was the original prototype used to generate all the category members (but was never presented during training) and was the prototype that was presented during the categorization test. This point will be critical in our later discussion.

Smith and Minda (2001) fitted the prototype model and the exemplar model by finding parameters (c and k_C) that minimized the sum of squared deviations between observed data and model predictions. The best-fitting predictions are shown in Fig. 10. As shown in the two left panels, the prototype model provided excellent accounts of the observed data for both controls and amnesics. By contrast, as shown in the two right panels, the exemplar model provided an exceedingly

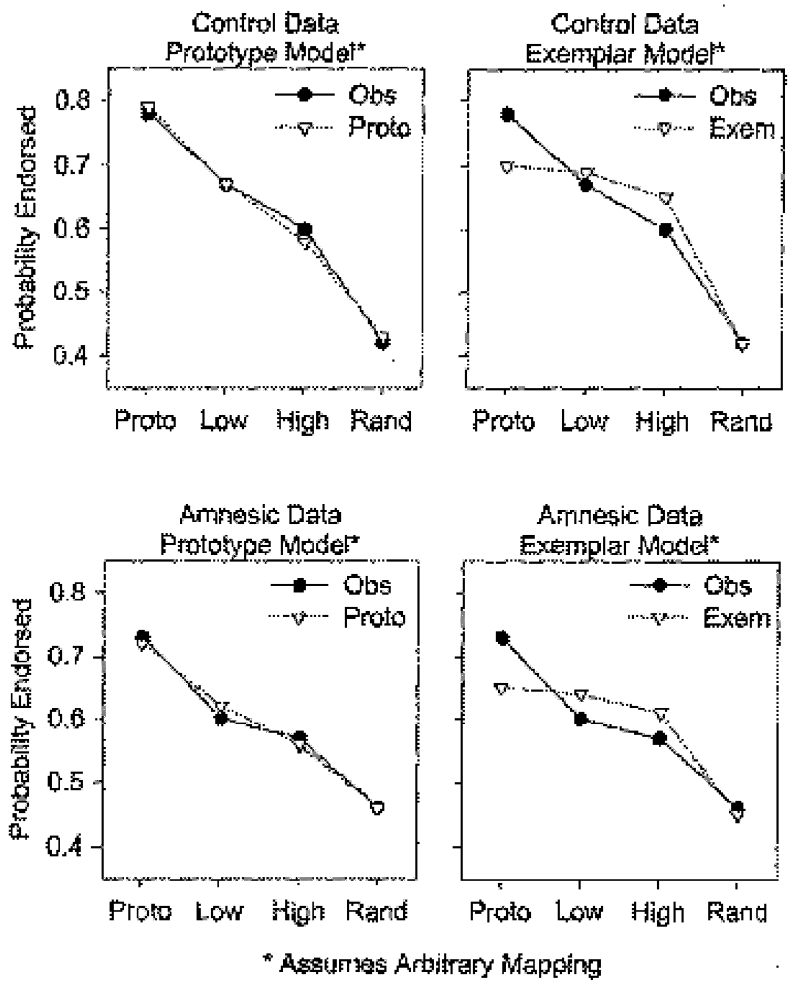

Fig. 10. Categorization predictions for a prototype model (left column) and an exemplar model (right column) for data from Knowlton and Squire (1993) for controls (top row) and amnesics (bottom row) from Smith and Minda (2001). The prototype model assumed a population-based prototype (see text). Simulations of both the prototype model (Proto) and the exemplar model (Exem) assumed arbitrary mappings in distance calculations involving random patterns (see text). Each panel displays the observed (Obs) and model predicted probability of endorsing prototypes (Proto), low distortions (Low), high distortions (High), and random patterns (Rand) as members of the studied category.

poor account of the observed data, quite unlike the fits reported by Nosofsky and Zaki (1998). Smith and Minda argued that the prototype model naturally predicts the steep typicality gradients from prototype to low distortions to high distortions, whereas the exemplar model was constrained to incorrectly predict a relatively flat typicality gradient. On these grounds, Smith and Minda claimed that the Knowlton and Squire (1993) results indeed provide support for a prototype-based, presumably implicit, categorization system that is entirely independent from an explicit, exemplar-based declarative memory system.

C. A Critique of the Critique

Nosofsky, Zaki, and Palmeri (2001) responded by pointing out a number of serious problems with Smith and Minda's analyses. Nosofsky et al. first questioned the use of gross physical measures of similarity in lieu of more psychologically valid measures of similarity, such as those obtained from subjective similarity ratings. Simply calculating the average distances between individual dots ignores any higher-order relational information (e.g., symmetry, density, coincidence) that would clearly play a major role in determining the similarity between dot patterns. Indeed, Palmeri and Nosofsky (2001) conducted experiments specifically aimed at demonstrating the importance of using psychological measures of similarity, such as multidimensional scaling, and not using physical measures of similarity, such as distances between dots. Specifically, Palmeri and Nosofsky showed that in some cases the physical central tendency of category exemplars (what would typically be considered the "category prototype") may be represented as an extreme point in psychological space rather than a central tendency in the psychological space. Although dot distances may capture some of the first-order similarities between patterns (e.g., that low distortions are more similar to a prototype than high distortions), they cannot be valid measures of the true psychological similarity between specific pairs of patterns.

Yet, even granting the use of dot distances as valid measures of psychological similarity, Nosofsky et al. (2001) also pointed out two serious flaws in the theoretical analyses reported by Smith and Minda (2001). First, recall that in the Knowlton and Squire (1993) dot pattern paradigm, the category prototype was a randomly generated dot pattern and the studied category members were 40 high-level distortions of the prototype. Prototype models generally assume that people learn categories by abstracting a prototype from the studied category members. This *sample-based prototype* (averaged across category members) is then used to classify new patterns during the subsequent categorization test. Yet, Smith and Minda erroneously assumed that the *population-based prototype,* which was originally used to generate the high-level distortions for the training session and which was presented for classification during the categorization test, but was never presented during training, was the enduring category representation. Although experience with an infinite number of distortions causes the sample-based prototype to converge onto the population-based prototype, the sample-based prototype is typically not identical to the population-based prototype, even with 40 training examples. In order to be theoretically sensible, the sample-based prototype should be assumed as the category representation because there is no way that a subject could discover the true population-based prototype given the limited number of category exemplars that were presented to them during training.

More critically, Nosofsky et al. (2001) noted that assuming the population-based prototype gives the prototype model an unfair advantage in accounting for the observed categorization responses. As shown in Fig. 10, one of the key empirical findings in the Knowlton and Squire (1993) experiments was a large prototype enhancement effect in which the prototypes were endorsed as category members over 10% more often than the low distortions. By assuming that the enduring category representation is the very same population-based prototype that is presented during the categorization test, the model is guaranteed to predict a large prototype enhancement effect since the presented prototype and the stored prototype are identical, and thus have a physical distance of zero. Again, there is no way for a human subject (nor a statistical learning algorithm) to induce the true population-based prototype given the examples that are experienced, so the large prototype enhancement effect predicted by this population-based prototype model cannot be based on any plausible psychological (or mathematical) principles.

Finally, and most critically, Nosofsky et al. (2001) also raised a serious concern with how Smith and Minda (2001) calculated physical distances between a critical subset of the dot patterns. For each pair of patterns, the Euclidean distance between each *corresponding dot* in the two patterns is computed, with the average distance serving as the measure of distance between the two patterns. But, what are the corresponding dots? For prototypes, low distortions, and high distortions, the correspondence problem is straightforward. The modeler knows which dot in a distortion corresponds with which dots in the prototype since the distortions were generated from the prototype. By extension, the modeler also knows which dot in one distortion corresponds with which dot in another distortion since both distortions were generated from the same prototype. But what about the random patterns? Consider the two dot patterns shown in Fig. 11, a high distortion of a prototype and a random pattern. Which dots in the random pattern correspond with which dots in the high distortion for purposes of calculating the distance between those patterns? Coming up with a solution to this correspondence problem is the key to any reasonable use of physical dot distances as surrogate measures of psychological similarity.

Smith and Minda (2001, p. 996) essentially argued that this apparently difficult correspondence problem was not a problem at all: "The distribution of logarithmic distance estimates is so narrow that any value from it would produce nearly identical modeling results. This means that ambiguity about dot correspondences has no in the middle of the distribution. Indeed, the mean of the distribution shown in Fig. 11 is 3.51 with a standard deviation of .15 and a 99% confidence interval ranging from 3.09 to 3.79. But, the minimum of the distribution (for the "optimal" correspondence) is just 2.60! To demonstrate the generality of this result, we generated high distortions from 1000 randomly generated prototypes and generated 1000 random patterns. The average *mean* distance between random patterns and high distortions across all possible correspondences for all 1000 pairs of patterns was 3.46, yet the average *minimum* distance was just 2.68. For comparison, across 1000 randomly generated stimulus sets, the average distances between the prototype, low distortions, and new high distortions and old high distortions were 2.16, 2.20, and 2.43, respectively (using the natural correspondence based on how the distortions were generated). Indeed, using an arbitrary correspondence, as used by Smith and Minda (2001), causes the exemplar model to predict a very flat typicality gradient because the resulting distance measures between random patterns and high distortions is so exorbitant (3.46 rather than 2.68). But using the much more sensible minimum-distance correspondence allows the exemplar model to instead predict a gradually rising typicality gradient because the distances (and hence similarities) between presented items and stored exemplars (high distortions) rise in an analogous manner. Thus, as further illustrated below, the main failure of the exemplar model in Smith and Minda's analyses does not stem from a fundamental failure of exemplar representations but rather from an improper use of arbitrary dot correspondences in calculating distances for a critical subset of the patterns.

D. AN EXEMPLAR-BASED INTERPRETATION REVISITED

To begin with, we first attempted to replicate the model fits using procedures similar to those originally used by Smith and Minda (2001). In addition to the original Knowlton and Squire (1993) experiment, Smith and Minda also fitted data from replications and extensions of this paradigm reported in a second experiment by Knowlton and Squire (1993) and in two experiments by Reber, Stark, and Squire (1998a,b). In this article, we also report model fits to Squire and Knowlton (1995), which was not included in the Smith and Minda analyses.[1] Table I displays summary fits of the population-based prototype model and the exemplar model using

[1] The test stimuli used by Nosofsky and Zaki (1998) and Palmeri and Flanery (1999) were identical to those originally used by Knowlton and Squire (1993) in that the very same set of dot patterns was used. The composition of test stimuli used by Reber et al. (1998a) was identical in that there were 4 repetitions of the prototype, 20 low distortions, 20 high distortions, and 40 random patterns, but each subject viewed a different set of randomly generated patterns. The second experiment of Knowlton and Squire (1993) provided only four unique training exemplars but the composition of types of test stimuli was identical to that used in their first experiment. The composition of test stimuli used by Reber et al. (1998b) was different from the others in that there were 4 repetitions of the prototype, 16 low distortions, 16 high distortions, and 36 new patterns (4 repetitions of a novel prototype, 16 low distortions of that prototype, and 16 high distortions of that prototype).

TABLE I

SUMMARY FITS OF POPULATION-BASED PROTOTYPE MODEL, SAMPLE-BASED PROTOTYPE MODEL, AND EXEMPLAR MODEL USING ARBITRARY MAPPINGS BETWEEN RANDOM PATTERNS AND CATEGORY REPRESENTATIONS

Source	Model	P	L	H	R	*c*	*k*	SSD
Knowlton & Squire (1993)	Observed	.78	.67	.60	.42			
(Controls)	Population prototype	.78	.69	.57	.44	.441	.285	.0017
	Sample-based prototype	.75	.71	.59	.42	.578	.191	.0027
	Exemplar	.72	.71	.64	.41	1.030	1.633	.0069
Knowlton & Squire (1993)	Observed	.73	.60	.57	.46			
(Amnesics)	Population-based prototype	.71	.64	.55	.46	.303	.414	.0024
	Sample-based prototype	.68	.65	.57	.45	.388	.319	.0051
	Exemplar	.66	.65	.60	.45	.680	4.628	.0088
Reber, Stark, & Squire (1998a)	Observed	.71	.55	.54	.40			
	Population-based prototype	.68	.60	.51	.41	.334	.465	.0049
	Sample-based prototype	.66	.62	.52	.40	.429	.348	.0084
	Exemplar	.63	.62	.56	.39	.780	4.215	.0120
Reber, Stark, & Squire (1998b)	Observed	.85	.66	.63	.40			
	Population-based prototype	.82	.72	.58	.42	.539	.219	.0077
	Sample-based prototype	.79	.74	.60	.40	.674	.145	.0115
	Exemplar	.75	.74	.66	.39	1.203	.951	.0171
Knowlton & Squire (1993)	Observed	.76	.71	.57	.31			
Experiment 2	Population-based prototype	.81	.69	.52	.35	.597	.240	.0063
	Sample-based prototype	.74	.72	.59	.30	1.111	.051	.0007
	Exemplar	.73	.71	.61	.30	1.461	.060	.0030
Squire & Knowlton (1995)	Observed	.70	.64	.58	.37			
(Controls)	Population-based prototype	.73	.64	.53	.41	.396	.375	.0043
	Sample-based prototype	.70	.66	.54	.39	.540	.251	.0016
	Exemplar	.67	.66	.59	.37	1.010	2.091	.0014
Squire & Knowlton (1995)	Observed	.70	.75	.59	.44			
(E.P.)	Population-based prototype	.76	.68	.57	.46	.383	.316	.0081
	Sample-based prototype	.74	.70	.59	.44	.521	.215	.0032
	Exemplar	.71	.70	.63	.43	.940	2.075	.0048

Note. Data and predictions are probability of endorsing each item type as a category member. P = Prototype, L = Low distortion, H = High distortion, R = Random pattern, *c* = sensitivity, *k* = response criterion, SSD = sum of squared deviations.

arbitrary mappings between random patterns and the category representations (prototypes or exemplars, respectively); for comparison, we also include fits of the sample-based prototype model using arbitrary mappings. In all but one case, the exemplar model fitted worse than either of the versions of the prototype model, as was shown by Smith and Minda. It may be instructive to note that the relative fits of the population-based versus sample-based prototype models was proportional to the degree of prototype enhancement observed in each particular experiment. For example, whereas the population-based prototype model better fitted the Knowlton and Squire (1993) results, which showed large prototype enhancement effects (see Nosofsky et al., 2001), the sample-based prototype model better fitted the Squire and Knowlton (1995) results, which showed far smaller or nonexistent prototype enhancement effects.

Next, we report fits of a sample-based prototype model and exemplar model using minimum-distance mappings between random patterns and category representations. Again, one of the issues that emerges when using average physical dot distance as a measure of similarity is how to solve the correspondence problem. One approach is to use the optimal correspondence that minimizes the distance between two patterns for every possible pair of patterns. One potential drawback of this approach is that the logical correspondence between different patterns generated from the same prototype may differ simply because the minimization criterion is enforced (Nosofsky et al., 2001). In addition, solving for the optimal minimal correspondence is very time consuming in that it requires 9! distance calculations for each pair of patterns. In some of the ensuing analyses, we generated predictions by averaging over 100 randomly generated stimulus sets—calculating all the necessary optimal distances between every tested item and every stored item, when simulating the exemplar model would require hundreds of *billions* of distance calculations, which is a practical hurdle in conducting the simulations.

Instead, Nosofsky et al. (2001) proposed a compromise solution for calculating distances when fitting the exemplar model. Before a random pattern is compared to the stored exemplars, the optimal correspondence between the random pattern and the sample-based prototype is first calculated. Then this particular correspondence is used in calculating the distances between the random pattern and each high distortion (the stored exemplars). We should emphasize that this method was not meant to imply in any way that prototypes are represented as part of the category—the sample-based prototype is only used to give a first approximation to an optimal correspondence between a random pattern and each of the high distortions rather than computing the optimal correspondence for each comparison individually. That is, only 9! distance calculations are necessary instead of $40 \times 9!$ distance calculations—as an important practical consideration, this simplifying assumption meant a difference of several days of simulation time rather than several months of simulation time.

Table II displays the summary fits for the sample-based prototype model and the exemplar model using the minimum-distance mappings described above. The fits to Knowlton and Squire (1993) are taken from Nosofsky et al. (2001). The predictions for the Knowlton and Squire study were generated by calculating the distances between the actual dot patterns used in those experiments; for the remaining studies, predictions were generated by averaging across 100 simulated sets of randomly generated dot patterns. As should be clear from Table II, when using a sample-based prototype rather than a population-based prototype and when using minimum-distance mappings rather than arbitrary mappings, the difference

TABLE II

SUMMARY FITS OF POPULATION-BASED PROTOTYPE MODEL, SAMPLE-BASED PROTOTYPE MODEL, AND EXEMPLAR MODEL USING MINIMUM-DISTANCE MAPPINGS BETWEEN RANDOM PATTERNS AND CATEGORY REPRESENTATIONS

Source	Model	P	L	H	R	*c*	*k*	SSD
Knowlton & Squire (1993)	Observed	.78	.67	.60	.42			
(Controls)	Sample-based prototype	.73	.72	.61	.41	2.048	.506	.0047
	Exemplar	.73	.71	.62	.41	1.298	.080	.0044
Knowlton & Squire (1993)	Observed	.73	.60	.57	.46			
(Amnesics)	Sample-based prototype	.67	.66	.58	.45	.862	.180	.0074
	Exemplar	.67	.65	.59	.45	1.376	2.022	.0071
Reber, Stark, & Squire (1998a)	Observed	.71	.55	.54	.40			
	Sample-based prototype	.67	.62	.49	.43	.590	.283	.0107
	Exemplar	.65	.63	.51	.41	1.619	.621	.0112
Reber, Stark, & Squire (1998b)	Observed	.85	.66	.63	.40			
	Sample-based prototype	.80	.74	.55	.45	.926	.104	.0187
	Exemplar	.79	.76	.58	.42	2.586	.040	.0168
Knowlton & Squire (1993)	Observed	.76	.71	.57	.31			
Experiment 2	Sample-based prototype	.77	.73	.50	.37	1.935	.011	.0092
	Exemplar	.74	.71	.52	.38	2.310	.010	.0084
Squire & Knowlton (1995)	Observed	.70	.64	.58	.37			
(Controls)	Sample-based prototype	.71	.66	.50	.43	.716	.203	.0095
	Exemplar	.70	.67	.52	.40	2.034	.205	.0046
Squire & Knowlton (1995)	Observed	.70	.75	.59	.44			
(E.P.)	Sample-based prototype	.75	.70	.54	.48	.730	.163	.0075
	Exemplar	.74	.72	.57	.45	2.064	.158	.0025

Note. Data and predictions are probability of endorsing each item type as a category member. P = Prototype, L = Low distortion, H = High distortion, R = Random pattern, *c* = sensitivity, *k* = response criterion, SSD = sum of squared deviations. Fits to Knowlton and Squire (1993) are from Nosofsky, Zaki, and Palmeri (2001).

between the prototype and exemplar models reported by Smith and Minda (2001) simply disappears (in fact, the exemplar model provides a numerically better fit in six of the seven datasets).

Nosofsky et al. (2001) did note that both the prototype model and the exemplar model predict far smaller prototype enhancement effects than are oftentimes observed in the experiments, and discussed some possible reasons for the elevated enhancement effects. Given the recent work presented at the beginning of this article, one plausible reason for the underpredicted prototype enhancement effect is that subjects may be learning about the category during the categorization test. In other words, the underlying category representation is not just a set of high distortions of the prototype acquired during initial learning, but rather includes the prototype (which is presented four times during the categorization test) and the numerous low distortions that are all very similar to the prototype.

Nosofsky et al. (2001) provided preliminary evidence that a simple enhanced version of the exemplar model (using minimum-distance mappings) in which category representations are augmented by information acquired during the categorization test does indeed predict a far larger prototype enhancement effect than the basic exemplar model based only on stored training exemplars. As shown in Fig. 12, this learning-during-transfer version of the exemplar model provided an excellent account of the Knowlton and Squire (1993) results. For comparison, the figure also shows a population-based version of the prototype model (using minimum-distance correspondences) that also fitted that data well (although we

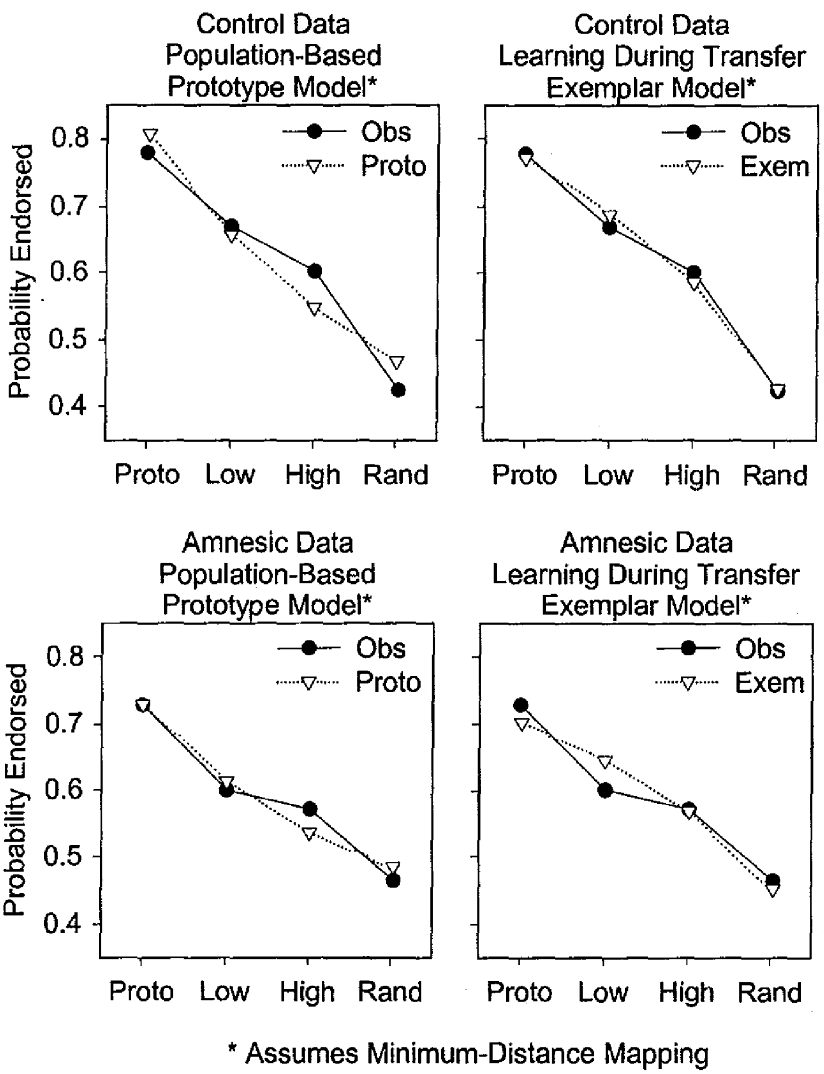

Fig. 12. Categorization predictions for a prototype model (left column) and an exemplar model (right column) for data from Knowlton and Squire (1993) for controls (top row) and amnesics (bottom row) from Nosofsky, Zaki, and Palmeri (2001). The prototype model (Proto) assumed a population-based prototype (see text). The exemplar model (Exem) assumed learning during transfer (see text). Simulations of both the prototype model and the exemplar model assumed minimum-distance mappings in distance calculations involving random patterns (see text). Each panel displays the observed (Obs) and model predicted probability of endorsing prototypes (Proto), low distortions (Low), high distortions (High), and random patterns (Rand) as members of the studied category.

still contend that using a population-based prototype is theoretically unfounded). The bottom-line result of these simulations is that the version of the dot pattern paradigm developed by Knowlton and Squire (1993) and used by other investigators cannot distinguish between prototype and exemplar representations, although other paradigms using dot pattern stimuli have indeed reported superior accounts by exemplar models over prototype models (e.g., Busemeyer, Dewey, & Medin, 1984; Palmeri & Nosofsky, 2001; Shin & Nosofsky, 1992).

Finally, as we noted earlier in this article, the extreme dissociation observed by Squire and Knowlton (1995) with the profound amnesic E.P. poses a clear challenge to the basic version of the exemplar model described by Nosofsky and Zaki (1998). To demonstrate this, we generated simulated predictions of categorization accuracy and recognition accuracy averaged over 100 randomly generated stimulus sets for values of the sensitivity parameter in a range of $.0 \leq c \leq 4.5$ in steps of .001.[2] For each value of c, we found values of k_C and k_R that produced unbiased responding (see Nosofsky & Zaki, 1998); in other words, these criteria produced an equal proportion of member/nonmember judgments in categorization and old/new judgments in recognition, respectively. Again, we used a minimum-distance mapping in all comparisons that involved a random pattern (i.e., between random patterns and high distortions in the categorization task and between all patterns in the recognition task). Figure 13 displays recognition accuracy plotted against categorization accuracy for all values of c in the simulated range. Within this range, as c increases, both recognition accuracy and categorization accuracy increase. However, it should be readily apparent that increases in c that produce large increases in recognition accuracy produce relatively modest increases in categorization accuracy. Indeed, it is this regularity which allowed the exemplar model to successfully account for the Knowlton and Squire (1993) dissociation as originally reported by Nosofsky and Zaki (1998). Yet, it should also be readily apparent that in order to predict *chance recognition,* the model is also forced to predict *chance categorization* as well. Thus, an exemplar model that bases responses on stored exemplars of the studied high distortions cannot account for the extreme dissociation reported by Squire and Knowlton (1995) for the profound amnesic E.P. These simulation results coupled with our experimental results argue for the acquisition of category information during the categorization test.

[2] With further increases in sensitivity, categorization accuracy eventually begins to fall, yet recognition accuracy remains at asymptote. Indeed, as c approaches infinity, categorization accuracy approaches chance (again) but recognition remains at perfect accuracy. In order to make sense of this prediction, it is important to remember that, in the Knowlton and Squire (1993) paradigm, during categorization, subjects are always tested on new items but that during recognition, subjects are tested on a combination of old and new items. With extremely large values of sensitivity, each memory trace becomes a completely unique entity (i.e., only a perfect match counts). This is ideal for recognition where the goal is to discriminate old from new items. But this is destructive to categorization in which new items are classified according to their similarity to old items; if only perfect matches count, then the generalization processes so crucial for categorization disappear entirely.

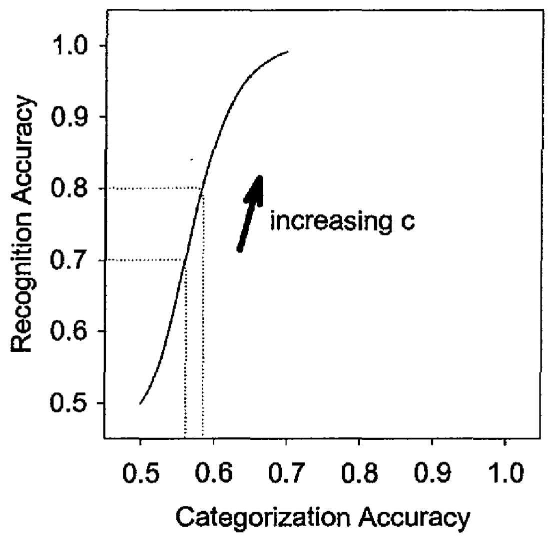

Fig. 13. Predictions of the exemplar model in the Knowlton and Squire (1993) paradigm for values of sensitivity (c) that vary from .0 to 4.5 with criteria (k_C or k_R) set to predict unbiased responding. The figure plots recognition accuracy against categorization accuracy across this range of parameters.

Smith and Minda (2001) did evaluate a version of the exemplar model that acquired category information during the categorization test in order to model the Palmeri and Flanery (1999) data. Not surprisingly, since their simulations were based on (arguably) erroneous distance calculation for random patterns, they reported similar failures of the exemplar model and instead reported successful accounts by a prototype model. One aim of some of our current research is to systematically investigate how well various learning-during-transfer versions of the exemplar model account for the results from Palmeri and Flanery (1999) that are summarized in the beginning section of this article. Although our initial investigations are showing that memory for only some of the test items is necessary to produce above-chance categorization, we must forgo discussion of these results for a future article.

IV. Other Experimental Paradigms

The dot-pattern paradigm is just one experimental procedure that has been used to contrast categorization and explicit memory by amnesics and normals. In this section, we review some other recent research showing dissociations between categorization and explicit memory, contrasting interpretations by multiple and single memory system accounts.

A. Learning Categories of Object-Like Stimuli with Discrete Features

Recent work by Reed, Squire, Patalano, Smith, & Jonides (1999) aimed to provide further evidence for multiple memory systems subserving categorization and explicit memory. Reed et al. generalized the investigation of preserved categorization in amnesia by using object-like stimuli with discrete features, which are quite unlike the continuously varying dot patterns used in the studies discussed earlier. The stimuli they used, which they called Peggles, were line drawings of animals that varied along nine binary-valued dimensions. As illustrated in Fig. 14, to create a category, a particular Peggle was designated as the prototype of the category. Category members were distortions of that prototype. Low distortions shared 7 or 8 features of the prototype, whereas high distortions shared only 1 or 2 features of the prototype. As an extreme, the antiprototype had all 9 features opposite to that of the prototype. Stimuli that shared 4 or 5 features of the prototype were designated neutral stimuli that were half way between the prototype and the antiprototype.

During an initial study phase, subjects viewed 40 low distortions of the prototype. Immediately after this initial exposure, subjects were told that the animals they just saw were all members of a category, called the Peggles, and were then asked to judge new animals as members or nonmembers of the Peggle category. During the test phase, subjects made member/nonmember judgments of 96 new

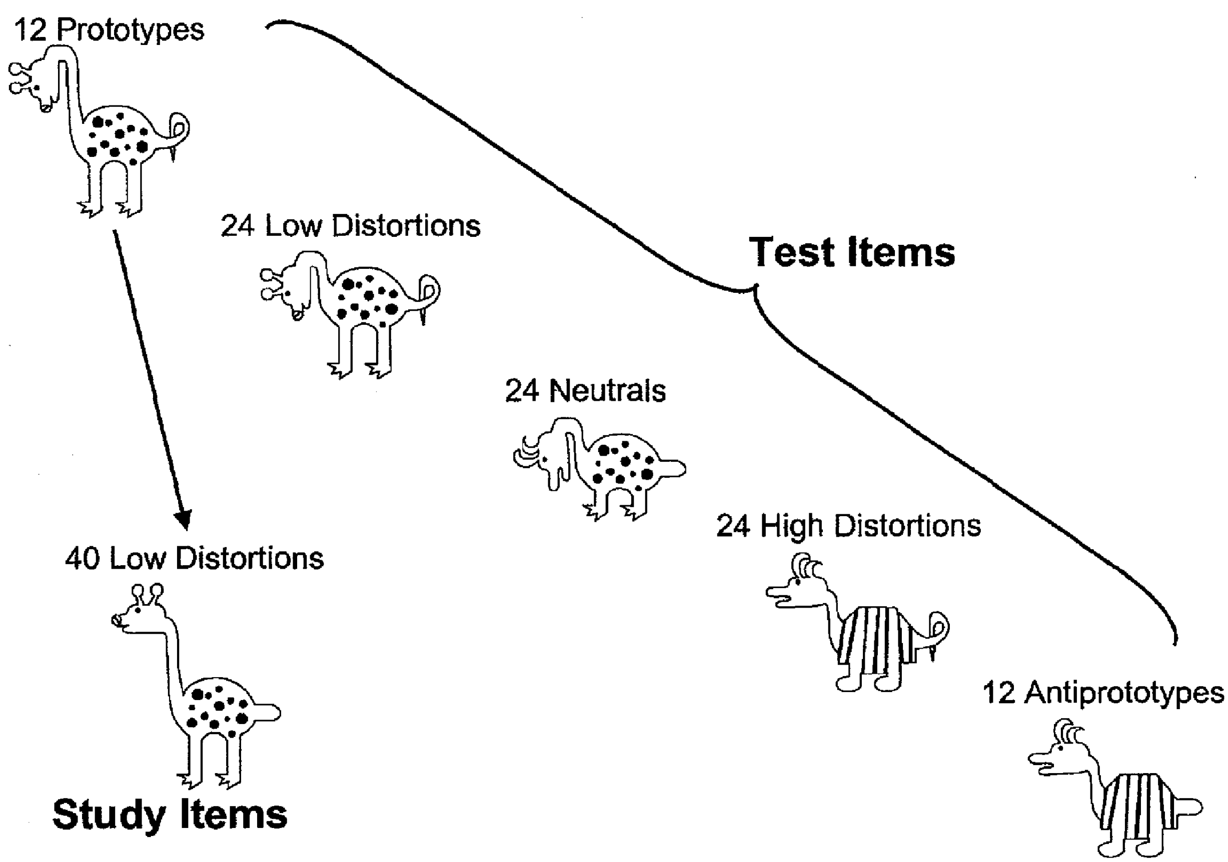

Fig. 14. Illustration of the Peggles used in Experiment 1 of Reed et al. (1999) that vary along 9 binary-valued dimensions. Study items were 40 low distortions of the prototype. Test items were 12 repetitions of the prototype, 24 low distortions, 24 neutral items, 24 high distortions, and 12 antiprototypes. Low distortions differed from the prototype along 1 or 2 dimensions, neutral items differed along 4 or 5 dimensions, high distortions differed along 7 or 8 dimensions, and the antiprototypes differed along all 9 dimensions.

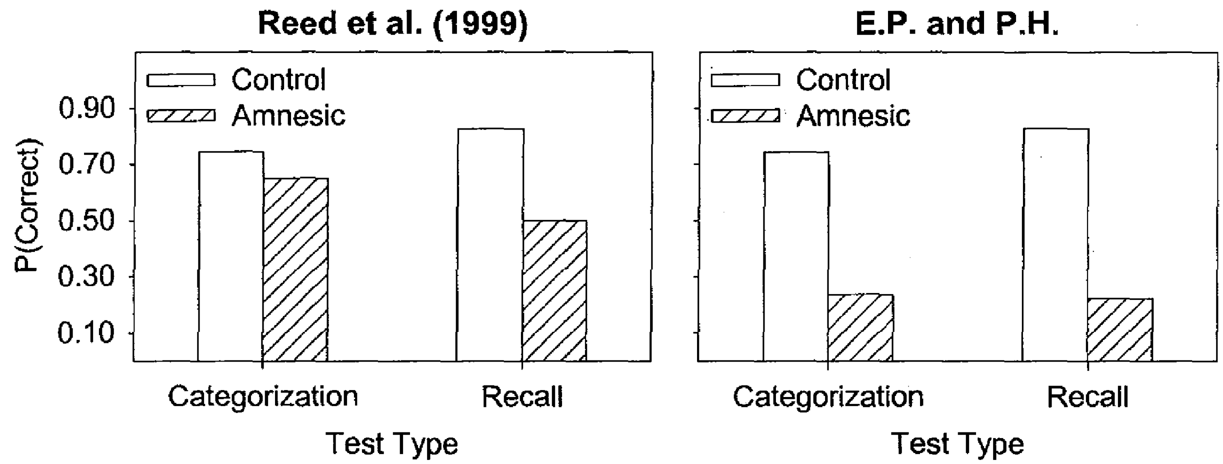

Fig. 15. Percentage correct categorization and cued recall performance for controls and amnesics from Experiment 1 of Reed et al. (1999). The left panel contains data from all amnesics and all normals tested. The right panel contains data for amnesics E.P. and P.H. compared to all normals.

stimuli presented one at a time without corrective feedback. The categorization test consisted of 12 repetitions of the prototype, 24 low distortions, 24 neutral stimuli, 24 high distortions, and 12 repetitions of the antiprototype. In addition, subjects were also tested on their ability to complete a cued recall test identifying the values of each of the 9 dimensions of the animals they had been shown.

As shown in the left panel of Fig. 15, Reed et al. (1999) found that amnesics were significantly impaired at explicit cued recall of the features of the animals but were not significantly impaired at categorizing the animals as Peggles or not.[3] Surprisingly, as shown in the right panel of Fig. 15, two of the amnesics actually categorized the test stimuli opposite to the way they should have (significantly *less* than chance?). That is, they mistakenly judged the prototype and low distortions to be *nonmembers* and mistakenly judged the antiprototype and high distortions to be *members,* and did so in a consistent fashion. Reed et al. suggested that amnesics had a spared implicit category learning system that had learned to partition members from nonmembers but that perhaps declarative memory was needed to explicitly remember which partition corresponded to the stimuli they had previously been exposed to (i.e., which partition corresponded to Peggles that were viewed earlier?).

1. Experiment 4: Learning about Categories during Testing

Following one of the themes of this article, we propose an alternative explanation. During the categorization test, subjects were shown the prototype many times (indeed, there were 12 repetitions of this single item) and were shown many low distortions that were very similar to the prototype. They were also shown the

[3] Note that for categorization, a P(correct) of .50 is considered chance since each item can be categorized as a member or nonmember, but for the open-ended cued recall test, "chance" performance is not defined in any similarly straightforward fashion.

antiprototype many times (indeed, there were 12 repetitions of this item as well) and were shown many high distortions that were very similar to the antiprototype. In other words, there were two clear clusters of stimuli presented during the categorization test, emphasized by the presence of the prototype or the antiprototype on 25% of the test trials. If subjects could discover the clear category structure embedded within the testing sequence so as to cluster stimuli into two groups, they would be able to partition the stimuli into two different categories. Obviously they would not be able to unambiguously decide which cluster corresponded to the category they were initially exposed to without relying on memory. Might this be a more reasonable explanation of the category switching by a subset of the amnesics in the Reed et al. study? And might this suggest that amnesics (and perhaps normals) may be relying more on information acquired during the categorization test than on information retrieved from long-term memory? By contrast, the explicit cued recall test cannot be completed without explicitly remembering what the cartoon animals looked like.

The goal of the following experiment was to test whether subjects might be categorizing in part by extracting information from the structure of the categorization test. Following Reed et al. (1999), our subjects initially studied 40 low distortions of a Peggle category prototype. After a varying delay, we then provided a cued recall test in which subjects described the feature values for all 9 dimensions of the animals.[4] Then subjects were given a categorization test. We randomly assigned subjects to one of three different conditions: Immediate, Delayed, and Novel. Subjects in the Immediate and Delayed conditions were tested in the same way as subjects in the Reed et al. (1999) experiment, except that subjects in the Immediate condition were tested immediately and subjects in the Delayed condition were tested 1 week later. As shown in Fig. 16, no significant difference in categorization was observed between the Immediate and Delayed groups (indeed, the Delayed group was numerically more accurate than the Immediate group), yet there was a significant difference in cued recall between the two groups.

One underlying motivation for the experimental design used in the Novel condition was the finding reported in Section II.D.4. Recall that in that experiment we had initially exposed some of the subjects to 40 high distortions of a prototype dot pattern and then tested subjects either on new stimuli generated from their studied prototype (Same condition) or on new stimuli generated from a novel prototype (Different condition). We observed no difference in categorization performance between these two conditions. The Novel condition in the present experiment had

[4] Unlike Reed et al. (1999), we gave subjects the cued recall test before the categorization test. This was necessary and entirely sensible because we were testing subjects after varying delays (i.e., if subjects tested after a 1-week delay were given the cued recall test after they completed the categorization test, their responses would be based on memory for what they just saw rather than memory for what they saw a week earlier).

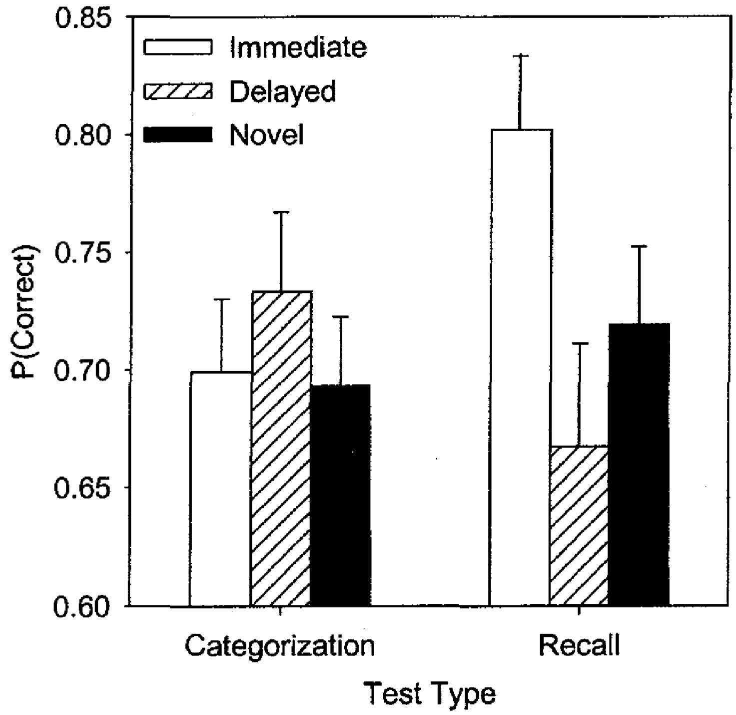

Fig. 16. Percentage correct categorization and cued recall performance on Peggles as a function of test condition in Experiment 4. We determined whether the prototype or antiprototype served as the defining member of each subject's internally defined category (see text). Correct categorization was then defined by either judging the prototype and low distortions as "members" and the antiprototype and high distortions as "nonmembers" or vice versa (see text). Correct cued recall was defined by the proportion of features recalled from each of the nine stimulus dimensions. In the Immediate Condition, subjects were tested within a single session. In the Delayed Condition, subjects were tested after one week. In the Novel Condition, subjects were tested after one week with a stimulus set defined by prototypes and antiprototypes that were neutral stimuli from the originally studied set. All subjects were tested on both categorization and cued recall.

a similar design in that subjects were initially exposed to distortions of a prototype, but when subjects were tested, the "members" and "nonmembers" were generated from a novel prototype.

Specifically, in the Novel condition, subjects also returned 1 week later to be given a categorization test. In this condition, the sequence of test stimuli contained an embedded category structure that actually contradicted what was presented during initial exposure. To do this, a neutral stimulus with respect to the prototype that was used to generate stimuli from the original exposure session was randomly selected and designated the "prototype" for purposes of creating a new categorization test sequence. From this novel prototype, low distortions, neutral stimuli, high distortions, and an antiprototype were created. Note that the "antiprototype" for this new structure would also be considered a neutral stimulus with respect to the prototype that was used to generate stimuli to which subjects had been originally exposed. The novel categorization test consisted of 12 repetitions of the novel prototype, 24 low distortions, 24 neutral stimuli, 24 high distortions, and 12 repetitions of the novel antiprototype.

Let us generate some predictions for the Novel condition. If subjects were categorizing based on what they had been previously exposed to, they should categorize the "prototype" and the "antiprototype" in this novel test sequence equally, as half way between the member and nonmember category with respect to what they had originally studied. However, if subjects were instead attending to the clear category structure embedded within this novel test sequence, they should group the "prototype" and its distortions in one category and group the "antiprototype" and its distortions in another category. Half of the subjects would call the "prototype" group the members and half would call the "antiprototype" group the members.

Scoring categorization performance for subjects in the Novel condition was somewhat more complicated than scoring in the other conditions. Essentially, what we first did was to measure the difference in membership endorsements for the "prototype" and the "antiprototype." Recall that if subjects were categorizing these two critical stimuli with respect to what they had studied, they should be indifferent at categorizing these stimuli as members or nonmembers. To the contrary, we found a 53.6% difference in membership endorsements for the "prototypes" and the "antiprototypes." Subjects were clearly discriminating between these stimuli when making category member judgments. Next, if a particular subject judged the "prototype" to be a member, then we scored categorizations of the low distortions as members and high distortions as nonmembers to be "correct" responses; on the other hand, if a particular subject judged the "antiprototype" to be a member, then we judged categorizations of the high distortions as members and low distortions as nonmembers to be "correct" responses. Figure 16 displays categorization accuracy for the Novel condition using this scoring method (actually, we scored the Immediate and Delayed conditions in the same way to make the reported results consistent across conditions). What should be clear from the figure is that subjects in the Novel condition discriminated between members and nonmembers in a way that was consistent with the structure embedded within the testing sequence and not on memory for what they had seen a week earlier. As with the dot pattern experiments reported earlier, we found comparable performance between subjects who were tested on categories they actually studied and subjects who were tested on categories that contradicted what they had actually studied.

2. *Summary*

In Experiment 4, we extended a paradigm used by Reed et al. (1999) to contrast categorization and recall by amnesics and normals. Reed et al. observed impairments in cued recall by amnesics compared to normals, but there was little difference in categorization between the two groups. However, they did observe that two of their amnesic individuals categorized members of the previously studied category as nonmembers and nonmembers as members. Although Reed et al.

interpreted these results in terms of an implicit memory for the category, we instead provided evidence that this ability to discriminate members from nonmembers might emerge from a clear category distinction embedded within the testing sequence. As we argued in the case of dot pattern categorization, perhaps the preserved ability of amnesics to categorize object-like stimuli with discrete features could be an artifact of the way the categorization tests were designed rather than evidence for independent memory systems subserving categorization and explicit memory.

B. Learning Categories Described by a Complex Quadratic Rule?

Filoteo, Maddox, and Davis (2001) investigated whether amnesics could learn to classify stimuli defined by a complex categorization rule. Adapting the well-known paradigm developed by Ashby and colleagues (e.g., Ashby & Gott, 1988; Ashby & Maddox, 1992), subjects learned two categories that were defined by multivariate normal distributions. In this paradigm, on every trial one of the two categories (normal distributions) is randomly selected and a stimulus is randomly sampled from that distribution. The subject classifies the stimulus as an A or a B and receives corrective feedback. Because any two normal distributions overlap, perfect performance is impossible in that an item that would otherwise be classified as a member of category A could have been selected from the tail of the category B distribution.

The categories used by Filoteo et al. (2001) were defined by the distribution parameters provided in Table III. Figure 17 displays contours of equal likelihood for the normal distributions that define the two categories. As shown in the figure, the two categories have a high degree of overlap. More importantly, learning to discriminate members of category A from members of category B requires integrating information from both dimension 1 and dimension 2. In the language of general recognition theory (Ashby & Gott, 1988; Ashby & Townsend, 1986), learning these categories requires the formation of a nonlinear, quadratic decision boundary that

TABLE III

Category Distribution Parameters from Filoteo, Maddox, and Davis (2001)

μ_1	μ_2	σ_1	σ_2	$\text{cov}_{1,2}$
Category A				
150	150	33	33	1052
Category B				
165	165	46	46	0

Note. μ_1 = mean along dimension 1, σ_1 = standard deviation along dimension 1, $\text{cov}_{1,2}$ = covariance of dimensions 1 and 2.

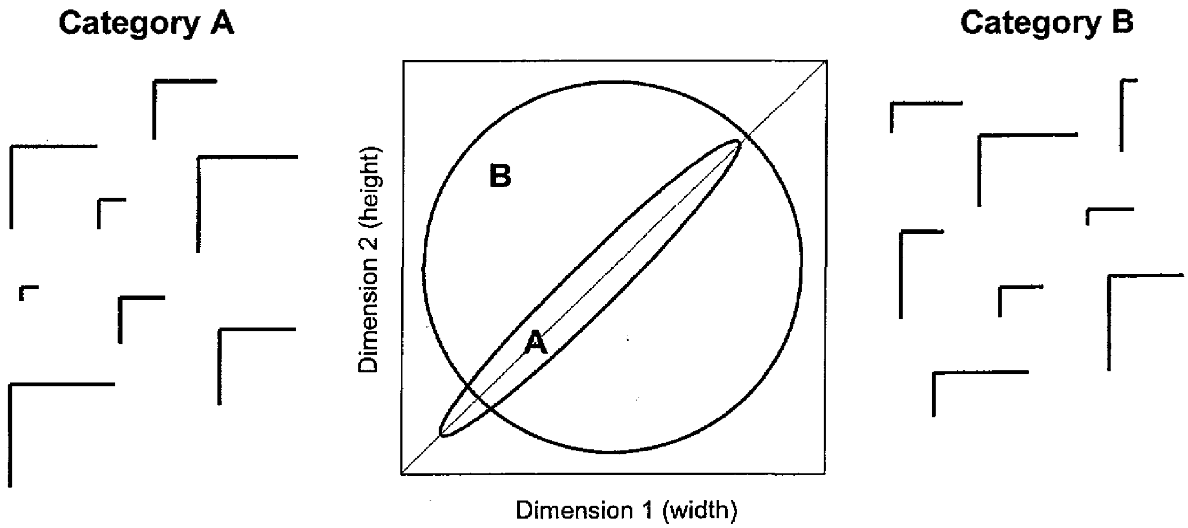

Fig. 17. The central panel displays the category structure used by Filoteo, Maddox, and Davis (2001). The ellipses give equal likelihood contours for the two multivariate normal distributions specified by the parameters in Table III. The dotted diagonal line specifies stimuli for which the value along dimension 1 equals the value along dimension 2. To the left and right of the central panel are illustrated examples of Category A ("squares") and examples of Category B ("rectangles").

combines information from both stimulus dimensions. This manipulation of the category structure was of particular theoretical importance because some recent work has suggested that amnesics cannot integrate information across multiple stimulus dimensions (Rickard & Grafman, 1998). This would imply that amnesics might be unable to learn categories defined by a quadratic decision boundary.

The physical stimuli used by Filoteo et al. (2001) consisted of a horizontal and a vertical line connected at the top left corner. The length of the horizontal and vertical lines varied in accordance with the category distributions shown in Fig. 17, and examples of each category are shown on the left and right hand sides of the figure. Note that the category A distribution consisted of stimuli for which the line lengths were highly correlated (i.e., given the parameters in Table III, the correlation between dimension 1 and dimension 2 was .966). In other words, as shown on the left side of the figure, the two line segments formed the left and top portions of a square (or a stimulus extremely similar to a square). Hence we will refer to category A as the "square" category. On the other hand, the category B distribution consisted of stimuli for which the line lengths were entirely uncorrelated. In other words, as shown in the right side of the figure, the two line segments formed the left and top portions of various rectangles. Hence we will refer to category B as the "rectangle" category. On each trial of the experiment, subjects were presented with a stimulus randomly drawn from either the square or the rectangle category, categorized it as a member of category A or category B, and received corrective feedback. Subjects completed six 100-trial blocks with an equal number of stimuli from each category presented per block.

Filoteo et al. (2001) observed the accuracy in the last block of 100 trials to be 85% for normals and 84% for amnesics. Overall, the learning curves for the amnesics and normals were virtually indistinguishable. They concluded that amnesics appear to be able to acquire categories defined by a complex quadratic rule. To test whether an amnesic could retain that rule over a period of time, they tested one amnesic and one normal after a 1-day delay. Subjects completed a single block of 100 trials in which they received corrective feedback on every trial, just as in the original training session. Accuracy was 92% for the normal individual and 89% for the amnesic. Thus, according to Filoteo et al., amnesics appear to be able to learn and retain a complex quadratic categorization rule, even though the amnesics scored in the bottom percentiles on a variety of standard clinical neuropsychological memory measures.

1. Experiment 5: Are Subjects Learning a Complex Quadratic Rule?

The Filoteo et al. (2001) results suggest that amnesics can learn and retain a category described by a complex quadratic rule that requires integrating information from two stimulus dimensions, height and width. Our first question was whether amnesics were truly learning an extremely difficult categorization rule, or whether this categorization problem might alternatively be described using a far simpler single-dimension rule. As noted earlier, Rickard and Grafman (1998) have shown that amnesics appear to retain the ability to learn simple unidimensional discriminations but are impaired at discriminations requiring an integration of multiple stimulus dimensions.

Indeed, the stimuli used by Filoteo et al. can also be described in an alternative way by rotating the dimensions by 45 degrees. As shown in Fig. 18, we can instead describe the stimulus dimensions in terms shape and area. Now, the square and rectangle categories vary along a single dimension and can be categorized by a very simple shape rule rather than a complex quadratic rule. Filoteo et al. rejected this possibility, arguing that their subjects were indeed learning a complex quadratic rule requiring integration of information along two independent stimulus dimensions. But, we are puzzled by how these subjects were able to learn a "complex" categorization rule so quickly, reaching asymptotic performance after less than 100 trials. Indeed, one of the amnesics was performing near asymptote after just 20 training trials. For comparison, other categorization experiments using multivariate normal distributions that appear to require the formation of a quadratic decision boundary may take normal subjects several days to reach asymptotic levels of performance (e.g., Ashby & Maddox, 1992). In addition, in recent work, Ashby et al. (1998) have argued that classifying such line segment stimuli may sometimes be accomplished using a simple verbalizable categorization rule using stimulus shape.

To illustrate that subjects may not be learning a complex quadratic rule, but instead may be learning a simple shape rule, we replicated and extended the

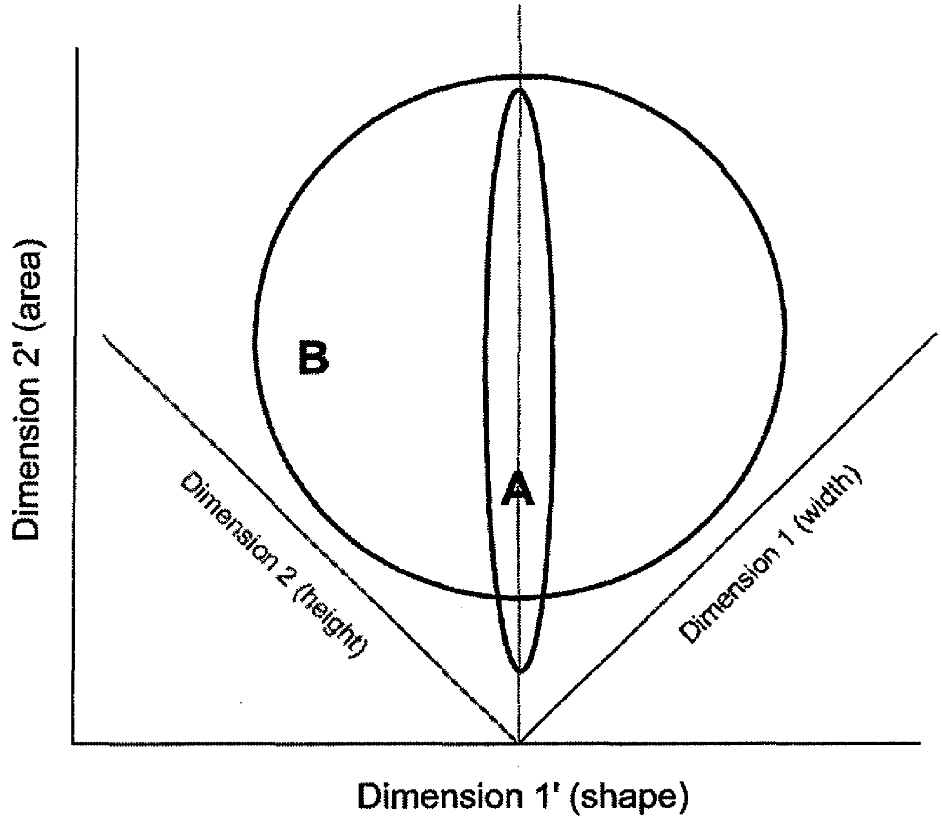

Fig. 18. A rotation of the dimensions in Fig. 17 by 45 degrees, yielding dimensions of shape and area.

Filoteo et al. (2001) study in the following way. In the first condition, we replicated their study using the same stimuli and category structures (Squares/Rectangles condition). In the second condition, subjects were trained on similar stimuli, but both multivariate category distributions shown in Fig. 17 were shifted along dimension 1 by 50 units. In this way, the category A distribution still had height and width highly correlated, but the values of height and width were not equal—in other words, the stimuli in category A were "squatty" rectangles of the same shape that varied in size and the stimuli in category B were other rectangles of varying shapes and sizes (Rectangle/Rectangle condition). In the third condition, we used very different stimulus dimensions of circles that varied in size containing an embedded diameter line that varied in orientation (Circle-Line/Circle-Line condition). Unlike the height and width of line segments, these two stimulus dimensions are incommensurable and cannot be readily integrated into any meaningful single dimension. Nor can a simple verbalizable rule be used to discriminate members of the two categories (see Ashby, Alfonso-Reese, Turken, & Waldron, 1998). Critically, in our experiment, the dimensions of the circle-line stimuli were roughly equated for discriminability with the height and width dimensions of the squares and rectangles (Maddox & Ashby, 1993). Five subjects completed each of the three conditions 1 week apart. The order of testing was Square/Rectangle, Circle-Line/Circle-Line, and Rectangle/Rectangle.

Filoteo et al. (2001) assumed that subjects were processing width and height of the line segments independently, such that they were required to learn a complex quadratic rule requiring the integration of information across two independent dimensions. If that were true, then the particular instantiation of dimension 1 and dimension 2 in the category structure shown in Fig. 17 should not matter at all. Thus, we might predict performance in the Circle-Line/Circle-Line condition to be comparable to performance in the Square/Rectangle condition. On the other hand, if subjects were instead using a simple shape rule in the Square/Rectangle condition, as we surmise, then performance in the Circle-Line/Circle-Line condition should be far worse. These stimuli indeed require integrating information across independent dimensions and require the formation of a complex decision rule.

The left half of Fig. 19 shows performance in the Square/Rectangle, Rectangle/Rectangle, and Circle-Line/Circle-Line conditions as a function of each block of 200 training trials (TB1, TB2, and TB3). Performance in the Square/Rectangle and Rectangle/Rectangle conditions were comparable (achieving 81 and 78% accuracy, respectively, in the final block). By contrast, performance in the Circle-Line/Circle-Line condition was terrible (58% accuracy). Indeed, all but one of the five subjects failed to exceed chance performance on the first block of trials, and three of the five subjects failed to exceed chance performance on the final block of trials. Our finding suggests that amnesics in the Filoteo et al. (2001) study may not have been learning a complex quadratic categorization rule at all,

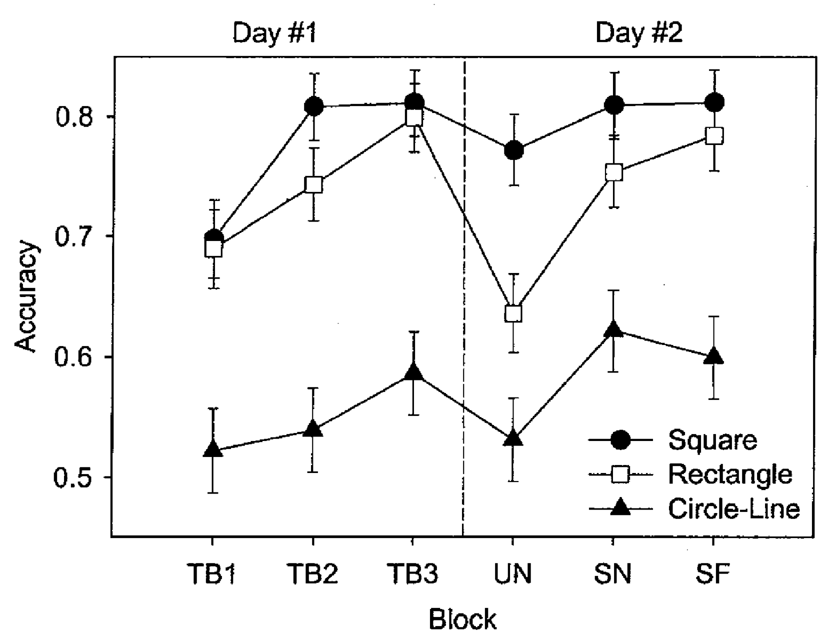

Fig. 19. Categorization accuracy in Experiment 5, our replication and extension of Filoteo et al. (2001). The left half of the figure shows performance on the three training blocks of 200 trials (TB1, TB2, and TB3) on Day 1 of each condition (Square/Rectangle, Rectangle/Rectangle, and Circle-Line/Circle-Line). The right half of the figure shows performance on the three types of transfer blocks of 200 trials on Day 2 of each condition. UN is the uniform condition with no feedback. SN is the structured condition with no feedback. SF is the structured condition with feedback.

but may have instead been learning a very simple shape rule. These limitations are important because Filoteo et al. instead argued that their results demonstrated intact categorization by amnesics under the most difficult of circumstances.

2. *Experiment 5: How Are Subjects Tested?*

Although we contend that amnesics in the Filoteo et al. (2001) study were learning a simple categorization rule, not a complex categorization rule, we are still impressed with the finding that one of their amnesic subjects was able to retain that categorization rule after a delay of an entire day. Specifically, during the first 10 trials of Day one, their amnesic was performing at chance. Yet, during the first 10 trials of Day two, their amnesic was performing at 80% accuracy. Certainly, one possibility is that amnesics can learn and retain simple categorization rules that do not require an integration of multiple stimulus dimensions (e.g., Rickard & Grafman, 1998). On top of this possibility, another possibility is that the way their subjects were tested for memory for the learned category is so unlike the way subjects are typically tested for explicit memory that making comparisons between categorization performance and explicit memory performance can be a precarious undertaking. Recall that one of the issues Nosofsky and Zaki (1998) emphasized, and that we demonstrated in simulations described earlier in this article, was that memory impairments can lead to significant deficits in explicit memory but only small deficits in categorization.

In Filoteo et al. (2001), one amnesic and 1 normal control returned after 1 day and were given the same categorization task they had been given on the first day. That is, on each trial, they saw a stimulus, classified it as an A or a B, and *received corrective feedback*. By contrast, on nearly every test of explicit memory ever conducted, subjects are never given corrective feedback, but instead are just asked to make a memory judgment which is then scored outside the presence of the subject. In the Filoteo et al. experiments, to what extent did their amnesic display an entirely unimpaired memory for categories or express a savings in relearning a very simple categorization rule? In addition, like the other experiments summarized in this article, Filoteo et al. sampled test stimuli from the very same distributions used to initially train subjects on the categories. That is, on half the trials they saw a square-like stimulus and on half the trials they saw a rectangle-like stimulus. Thus, like the experiments discussed previously in this article, the structure of the testing sequence served as a further cue to inform the subject how the studied categories were structured. In order to make categorization and explicit memory tests as comparable as possible, it is necessary to remove the category structure and to remove the corrective feedback.

To show that different kinds of categorization tests can reveal different levels of memory for studied categories, we brought our subjects back after 1 day and tested them in three different ways. Each test block consisted of 200 trials. First, we tested subjects *without feedback* on stimuli drawn from a *uniform distribution* across the set of possible stimuli (stimuli were sampled from a 5 × 5 grid that spanned the space in which most stimuli were selected)—uniform structure without feedback is indicated by UN in Fig. 19. Second, we tested subjects *without feedback* on stimuli drawn randomly from the two *category distributions*—category structure without feedback is indicated by SN in Fig. 19. Third, we *retrained subjects with feedback,* as was done by Filoteo et al. (2001)—category structure with feedback is indicated by SF in Fig. 19.

Although subjects reached comparable levels of performance in the Square/Rectangle and Rectangle/Rectangle condition by the end of the first day, as shown in Fig. 19, subjects were significantly better when tested on the uniform distribution without feedback (UN) in the Square/Rectangle condition than the Rectangle/Rectangle condition. By contrast, in the other two testing conditions (without feedback and with feedback), performance was comparable between the two conditions. Thus, different kinds of categorization tests can reveal very different levels of memory for previously learned categories. Just examining performance in the structured test with feedback would have led to the erroneous conclusion that subjects retained information about the category structures in the two different conditions equally well. By extension, we again argue that when testing amnesics on categorization and recognition, it is critical that the two tests be equated as much as possible: Remove corrective feedback and remove the informative category structure.

C. Categorization and Recognition in Artificial Grammar Learning

Finally, another experimental paradigm in which researchers have investigated dissociations between categorization and explicit memory is artificial grammar learning (e.g., Knowlton, Ramus, & Squire, 1992; Knowlton & Squire, 1994, 1996). In an artificial grammar learning experiment (Reber, 1969, 1989), subjects study letter strings that are generated from a finite state grammar. An example grammar is shown in Fig. 20. To generate a "grammatical" letter string, start at the left side of the network (IN) and follow the arrows until an exit point is reached (OUT). For each arrow that is followed, append the letter associated with the arrow to the letter string. For example, for the grammar shown in Fig. 20, the strings LCRRM, MTCCM, and MTTTL would be grammatical in that they follow the rules of the grammar, but the strings LCCL, MRLLT, and MTCTTM would be ungrammatical. In a typical artificial grammar learning experiment, subjects first memorize a set of letter strings that are generated from the finite state grammar. In a categorization task, subjects are then told that the letter strings they memorized were all generated by a complicated set of rules and are asked to judge new strings as grammatical or ungrammatical (half are generated from the grammar and half are not). For comparison, in a recognition memory task, subjects are instead asked to discriminate between old and new letter strings.

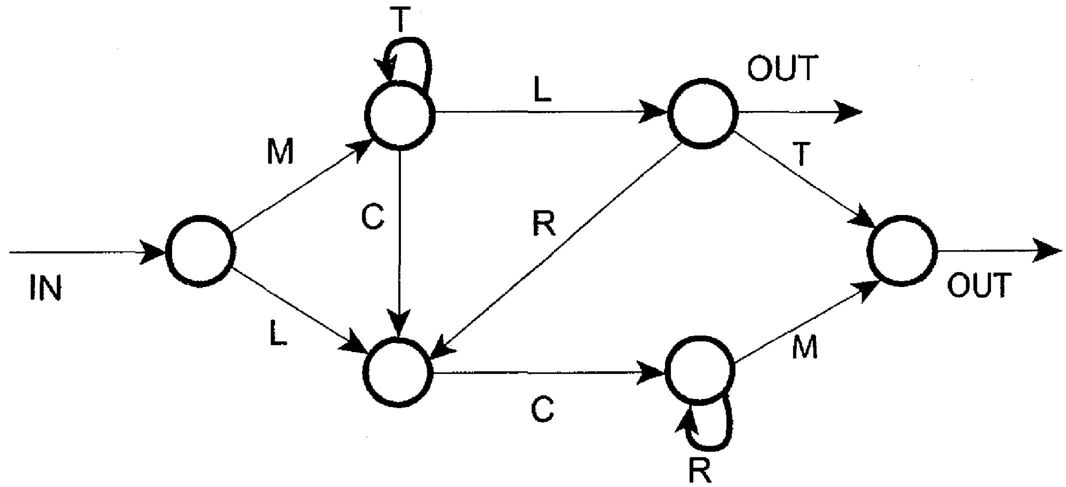

Fig. 20. An example of a network specifying an artificial grammar. The network is entered at the left (IN). Letter strings are generated by following the arrows through the network until an exit arrow is reached (OUT). When multiple arrows leave a node, each arrow has an equal probability of being followed. When a particular arrow is followed, the letter associated with that arrow is appended to the letter string. (From Annette Kinder and David Shanks, 'Amnesia and the Declarative/Nondeclarative Distinction: A Recurrent Network Model of Classification, Recognition, and Repetition Priming,' *Journal of Cognitive Neuroscience,* 13:5 (September, 2001), pp. 648–669. © 2001 by the Massachusetts Institute of Technology.

As shown in upper panel of Fig. 21, Knowlton et al. (1992) found that amnesics were not significantly impaired at categorizing letter strings as grammatical or ungrammatical, but were significantly impaired at recognizing letters strings as old or new (see also Knowlton & Squire, 1994, 1996). As with the other cases discussed in this article, this dissociation was taken as further evidence for independent systems governing categorization and recognition memory.

However, as with Nosofsky and Zaki (1998), Kinder and Shanks (2001) recently provided an alternative single-system explanation for the dissociation between categorization and recognition in artificial grammar learning. To do so, they adapted a successful connectionist model of artificial grammar learning based on a simple recurrent network (Cleeremans & McClelland, 1991; Elman, 1990). As with Nosofsky and Zaki (1998), Kinder and Shanks assumed that the difference between amnesics and normal controls was manifest in the change of a single parameter of the model, namely the learning rate of the connectionist network. Specifically, amnesics were assumed to have a lower learning rate than normals. As a single-system model, both categorization and recognition judgments were mediated by the same network, but because the stimulus sets for categorization and recognition were different, quite different predictions could emerge for the two tasks. As shown in the lower panel of Fig. 21, this simple recurrent network with different learning rates was able to account for the Knowlton et al. (1992) data quite well. As with Nosofsky and Zaki (1998), and as we showed earlier, a parameter difference can cause a relatively small difference in categorization but can lead to a relatively large difference in recognition. Although the modeling framework is quite

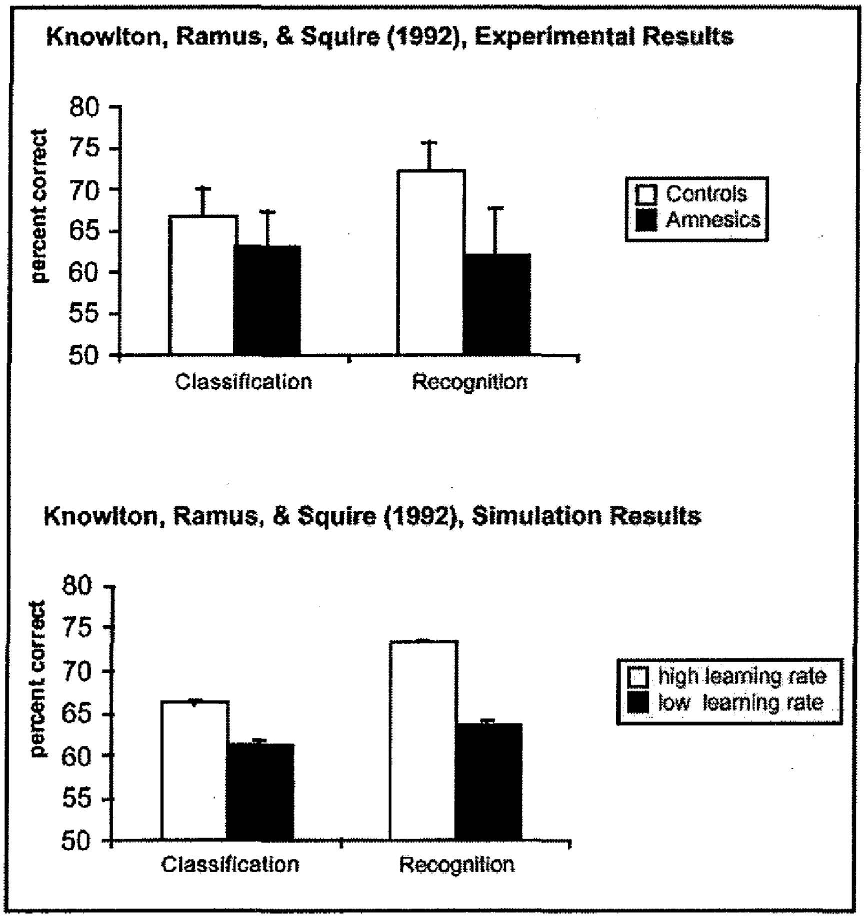

Fig. 21. Top panel shows observed classification and recognition accuracy by amnesics and normal controls from Knowlton et al. (1992). Bottom panel shows predicted classification and recognition accuracy by SRN for low learning rate (simulating amnesics) and high learning rate (simulating normal controls) from Kinder and Shanks. (From Annette Kinder and David Shanks, 'Amnesia and the Declarative/Nondeclarative Distinction: A Recurrent Network Model of Classification, Recognition, and Repetition Priming,' *Journal of Cognitive Neuroscience,* 13:5 (September, 2001), pp. 648–669. © 2001 by the Massachusetts Institute of Technology.

different (simple recurrent networks instead of exemplar models), the simulations by Kinder and Shanks (2001) provide further evidence that behavioral dissociations that seem to suggest multiple independent systems can often be explained by the operation of a single system suitably impaired to simulate brain damage.

V. Final Thoughts

Are dissociations between categorization and explicit memory evidence for independent memory systems? In experiments using dot patterns, amnesics appear to categorize as well as normals, but are significantly impaired at recognition memory (Knowlton & Squire, 1993; Squire & Knowlton, 1995). In experiments using stimuli with discrete features, amnesics appear to categorize as well as normals, but are significantly impaired at cued recall (Reed et al., 1998). In experiments using categories defined by multivariate normal distributions, amnesics appear to categorize as well as normals, but are significantly impaired on neuropsychological tests of explicit memory (Filoteo et al., 2001). And in artificial grammar learning experiments, amnesics appear to make grammaticality judgments as well as normals, but are significantly impaired at recognition memory (Knowlton et al., 1992; Knowlton & Squire, 1994, 1996). In addressing the implications of these dissociations, this article described two lines of research aimed at understanding why amnesics appear to show preserved memory for categories yet show impaired explicit memory for other kinds of information without needing to posit independent memory systems.

A. Findings From Computational Modeling

Computational models of human cognition aim to instantiate psychological principles involved in representing information in the environment, retrieving information from memory, storing information and creating new representations, utilizing information to make decisions, and so forth, in terms of well-specified computational and mathematical formalisms. By specifying a theory in this level of detail, it is then possible to test specific predictions of the theory, making it possible to falsify the theory. A typical approach to testing a computational model is to find values of the free parameters of the model that minimize the deviations between the model predictions and the observed data. The number of free parameters is a reasonable first approximation to the inherent flexibility of a model in accounting for particular patterns of observed data. The ideal psychological model would have zero free parameters, in which case the model would perfectly predict observed behavior a priori, a situation perhaps best approximated by certain physical laws of motion. In the other extreme, a model could have so many free parameters that it could account for any possible pattern of observed results, making the model entirely unfalsifiable.

When comparing models, it is necessary to equate the models for the number of free parameters, or to use fit statistics that penalize a model for the number of free parameters, or to use fit statistics that penalize a model for how flexible it is (Myung, 2000). A model with more free parameters and more flexibility is favored over a simpler model only if the more complex model provides a significantly better account of the observed data even after the various penalities for additional parameters and flexibility are imposed. In other words, the approach is to start with the simplest and most parsimonious model possible and to add complexity only when necessary. This is similar to the approach that verbal theory development often takes, but with a statistical underpinning for deciding when additional complexity is warranted.

Of particular relevance to the present discussion are cases where a simple model is a special case of a more complex model. That is, by restricting a subset of the parameters of the more complex model, the simple model emerges mathematically (e.g., a linear regression function is a special case of a quadratic regression function in that a linear function can be derived from a quadratic function by setting the constant for the quadratic term equal to zero). In terms of model testing, when two models have such a hierarchical arrangement, the more complex model is guaranteed to account for observed data better than the simple model that it contains as a special case (e.g., a quadratic regression function always provides a better fit to observed data than a linear regression function). Under these circumstances, it becomes critical to examine the improved fit of the more complex model using statistical criteria instead of absolute fit measures.

In this regard, some instantiations of a multiple memory systems theory could be viewed as containing a single memory system theory as a special case. For example, one could propose an exemplar model of explicit memory and an independent exemplar model of categorization, each with their own unique representations and their own unique free parameters; this multiple memory system theory is certainly viable, but it seems prudent to first consider a single system theory where categorization and explicit memory share parameters and representations. Alternatively, Knowlton and Squire (1993) and Smith and Minda (2001) proposed an exemplar model of explicit memory and an independent prototype model of categorization, each with their own unique representations and their own unique free parameters (and presumably with their own unique neural instantiation); again, although certainly viable, the modeling results we summarized in this article do not seem to warrant this additional theoretical complexity (Nosofsky & Zaki, 1998, 1999; Nosofsky et al., 2001). The language of research proposing single system accounts (e.g., Nosofsky & Zaki, 1998; Kinder & Shanks, 2001) does not suggest a single system model providing a *superior* account of observed data than a multiple system model—such a claim would be unfounded given the potential hierarchical arrangement of single system models within multiple system models. Instead, researchers note that "a single-system model is sufficient to explain categorization and recognition of stimuli generated by an artificial grammar . . ." (Kinder & Shanks, p. 15), or that "the single-system exemplar model provides an equally viable account of the categorization-recognition dissociation as do the multiple-system approaches . . ." (Nosofsky et al.), or that ". . . various of the important dissociations are also apparently consistent with the idea that a single exemplar-based memory system underlies categorization and recognition, as long as one allows for plausible differences in parameter settings across groups" (Nosofsky & Zaki, 1998, p. 255).

Indeed, one of the concerns with multiple memory system accounts is the potential proliferation of independent memory systems (see Roediger, Buckner, & McDermott, 1999). One piece of evidence for independent systems subserving categorization and recognition is the dissociation observed in amnesic subjects in the dot pattern and artificial grammar learning tasks (Knowlton & Squire, 1992, 1993). But, simple behavioral dissociations are notoriously weak evidence for independent systems, for a variety of reasons (see Hintzman, 1990). Instead, the "gold standard" for independent systems in neuropsychological research is the double dissociation, whereby patient A can do task 1 but not task 2, but patient B can do task 2 but not task 1. Such a double dissociation emerged with results reported by Knowlton, Mangels, and Squire (1996) in that Parkinson's Disease (PD) patients were significantly impaired at learning a probabilistic classification task but could explicitly remember aspects of the task, and amnesics patients were unimpaired at learning a probabilistic classification task but were significantly impaired at explicit memory for the task. However, closer examination of the data makes the true description of the double dissociation somewhat murky. Although PD patients were significantly impaired early in learning, they eventually reached the performance level of amnesics after 100 trials. However, both PD patients and amnesics were significantly worse than normal controls at this later stage in learning. This result is troubling because, by an independent systems account, amnesics should be performing just as well as normal controls throughout the task. So, what is the explanation for impaired performance by amnesics in this classification task? "Continued training may allow information to become available from declarative memory, that is, the controls and the PD patients may have eventually detected and memorized some of the cue–outcome associations" (Knowlton et al. 1996, p. 1401). Although certainly a viable explanation, such theoretical accounts risk becoming eminently unfalsifiable when any, perhaps unexpected, deficit by amnesics is simply explained by the amnesics' lack of declarative memory.

Putting that criticism aside, let us accept for now the double dissociation where amnesics can categorize but not recognize and PD patients can recognize but not categorize. By this account, we might predict that PD patients should also show a deficit on dot pattern classification and artificial grammar learning, both of which are classification tasks and on both of which amnesics have been reported to perform as well as normals. Reber and Squire (1999) tested PD patients on dot pattern classification and artificial grammar learning. Perhaps surprisingly, PD patients performed entirely normally on both tasks. But rather than reevaluating the original independent memory systems account, this finding provided evidence for yet another independent memory system. Recognition memory is served by a declarative memory system mediated by the hippocampal formation; dot pattern classification learning is served by a perceptual learning system mediated by neocortex; and probabilistic classification learning is served by procedural habit learning system mediated by the striatum of the basal ganglia. Perhaps. But there may be more parsimonious explanations that should be considered first.

Without question, double dissociations are more compelling evidence than simple dissociations. Clearly one interpretation of a double dissociation is that there are independent systems mediating the two tasks. However, another, equally viable, interpretation is that there are critical differences in the component processes that make up the two tasks (e.g., Moscovitch, 1992, 1994; Roediger et al., 1999). For example, Nosofsky and Zaki (1998) proposed a single system account in which amnesia led to impairments in memory sensitivity but PD led to impairments in response selection. With the appropriate change in a single parameter of the model, either to memory sensitivity or to response determinism, Nosofsky and Zaki (1998) were able to account for deficits by amnesics at recognition and deficits by PD patients at categorization within a single system. Similarly, Kinder and Shanks (2001) simulated a double dissociation between perceptual priming and recognition memory by varying single parameters associated with different aspects of their computational model. While double dissociations may suggest some structural organization of a system—into modules for memory and for response selection in Nosofsky and Zaki (1998) or into modules for memory and for perceptual processing in Kinder and Shanks—they do not necessarily compel functionally independent memory systems.

For example, Fig. 22 shows a depiction of the ALCOVE model (Kruschke, 1992), a connectionist version of an exemplar model of categorization. Each input node represents the value of an input stimulus along a particular psychological dimension. Each dimension is weighted by a learned selective attention gate. The presented stimulus activates exemplar memory nodes according to the similarity between the presented stimulus and the exemplar. Exemplars activate category output nodes along weighted connections that are learned. This is a single system model with a number of identifiable components. Each of these components could have a separate neural representation that could be selectively impaired by localized brain damage. Even "memory" in the network has multiple instantiations: there is memory for the exemplars in the hidden layer, but there is also memory for

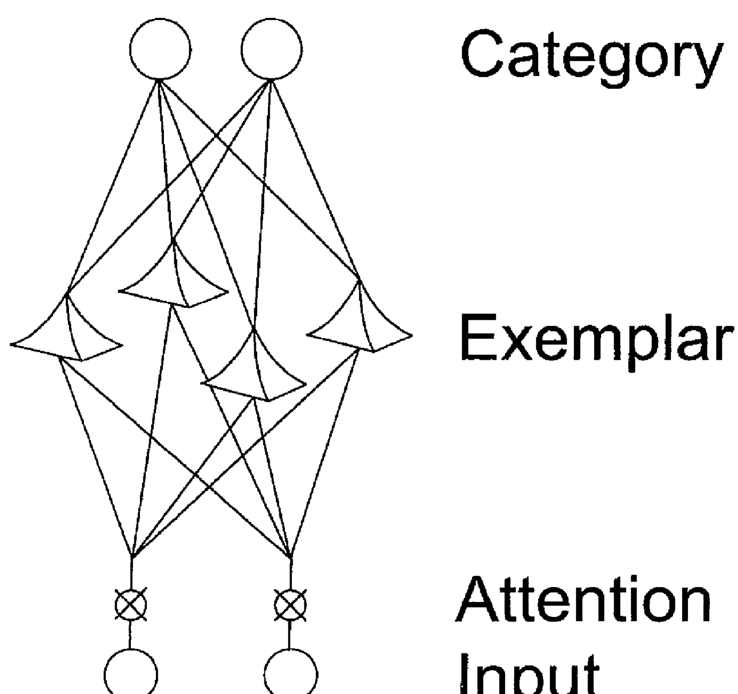

Fig. 22. The architecture of the ALCOVE model. (From Kruschke, 1992.)

the connections between exemplars and categories (and further, there is memory for particular patterns of selective attention to stimulus dimensions). Both kinds of memory would be necessary for categorization, but perhaps only one kind of memory would be necessary for recognition (that based on the exemplar nodes). These different kinds of memory do not seem to fit the standard operational definition of a memory system in that they are highly interactive within a unified processing architecture. Selectively impairing different aspects of the ALCOVE network can lead to a variety of different qualitative impairments. While dissociations and double dissociations may indeed dictate a modular (or semimodular) organization, they do not necessarily dictate independent systems with their own unique representations and processes.

In general, whereas the proliferation of multiple systems can be a natural consequence of a simplistic neuropsychological interpretation of behavioral dissociations and double dissociations, computational modeling approaches are far more conservative in positing separate systems. Indeed, in the area of perceptual categorization, there is currently a great deal of debate over the purported existence of separate rule-based and exemplar-based systems. Several recent computational models have proposed separate rule-based and exemplar-based (or otherwise implicit) subsystems (e.g., Ashby et al., 1998; Erickson & Kruschke, 1998; Palmeri, 1997) or a mixture of rule-based and exemplar-based representations (e.g., Anderson & Betz, 2001, Love, Medin & Gureckis, in press; Vandierendonck, 1995), yet the need for positing separate systems is still under serious debate (e.g., see Johansen & Palmeri, in press; Nosofsky & Johansen, 2000).

B. Tasks Used to Study Categorization and Explicit Memory

In most categorization paradigms, subjects acquire information about novel categories during an initial study phase and are later tested on the knowledge they have acquired about those categories. Patterns of observed responses during the categorization test can serve as a window on the types of memory representations that are formed about a category, be they prototypes (e.g., Posner & Keele, 1968), rules (e.g., Nosofsky, Palmeri, & McKinley, 1994), exemplars (e.g., Medin & Schaffer, 1978), decision boundaries (e.g., Ashby & Gott, 1988), or some combination of these (e.g., Ashby et al., 1998; Erickson & Kruschke, 1998; Johansen & Palmeri, in press). An even more fundamental question is whether or not subjects have acquired any information about the categories during the study phase. As we have discussed, this is particularly important when investigating whether certain brain-damaged individuals, such as amnesics and PD patients, can perform categorization tasks.

In explicit memory experiments, subjects study a set of items and are later tested on their memory for those items with recognition or recall tests. In most cases, performance on the memory test is entirely a function of information acquired during the study phase of the experiment. Similarly, in categorization experiments, it is typically assumed that performance on a categorization test only reflects knowledge acquired about the category during the initial study task. The particular choice of test stimuli may reflect the specific hypotheses that are being evaluated within that study—subjects might be tested on a previously unseen prototype used to generate the studied category examples (e.g., Posner & Keele, 1968), or they might be tested on very extreme category examples (e.g., Nosofsky, 1991), or they might be tested on new stimuli that prove diagnostic with respect to certain theoretical alternatives (e.g., Nosofsky & Palmeri, 1997)—but it is generally assumed that the particular choice of test stimuli will not influence subjects' apparent knowledge of the previously acquired categories in any systematic way.

However, there are reasons to question this assumption under certain conditions. For example, as a test of the generalized context model, Nosofsky (1986) trained two subjects on a variety of category structures. For each structure, subjects studied instances of two categories with feedback and were then tested on the old stimuli and new stimuli. In order to increase the statistical power in examining individual subject data, Nosofsky tested each individual many times (approximately 3500 trials for each category condition). When theoretically modeling the results from this particular set of studies with the GCM, Nosofsky found that it was necessary to augment the exemplar model by assuming that the new transfer stimuli, which were presented many times, became an integral part of the stored category representations.

In studies of amnesics and normals, it has been implicitly assumed that knowledge expressed on the categorization test or the explicit memory test reflects information acquired during the initial study phase of the experiment. Our recent results summarized in this article (Palmeri & Flanery, 1999) provide evidence that a great deal of information about categories may instead be acquired through exposure to a categorization test task, in the absence of any prior study and sometimes in opposition to prior study. As such, it is possible that explicit memory tasks used to test amnesics and normals may be true tests of memory, but some categorization tasks may reflect both long-term memory for previously studied information and information acquired more recently during the categorization test itself. Our results, as well as other recent findings (Buchner & Wippich, 2000), strongly argue for the absolute necessity of equating different tasks as much as possible, particularly when the aim is to document whether particular classes of brain-damaged individuals can perform some tasks, but not others.

ACKNOWLEDGMENTS

This work was supported by NIMH Grant R01 MH61370 and NSF Grant BCS-9910756. We thank Shannon Bader, Allison Bell, Lyndsay Boggess, Jeff Burgess, Karinne Damadio, Talia Valdez, Robert Emms, Kinna Patel, and Brooke Schaper for testing participants in these experiments. We also thank Gordon Logan, Brian Ross, and David Shanks for comments on early versions of this article.

Correspondence should be addressed to Thomas J. Palmeri, Department of Psychology, 301 Wilson Hall, Vanderbilt University, Nashville, TN 37240, thomas.j.palmeri@vanderbilt.edu.

REFERENCES

Anderson, J. R., & Betz, J. (2001). A hybrid model of categorization. *Psychonomic Bulletin & Review, 8,* 629–647.

Ashby, F. G., Alfonso-Reese, L. A., Turken, A. U., & Waldron, E. M. (1998). A formal neuropsychological theory of multiple systems in category learning. *Psychological Review, 105,* 442–481.

Ashby, F. G., & Gott, R. E. (1988). Decision rules in the perception and categorization of multidimensional stimuli. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 14,* 33–53.

Ashby, F. G., & Maddox, W. T. (1992). Complex decision rules in categorization: Contrasting novice and experienced performance. *Journal of Experimental Psychology: Human Perception and Performance, 18,* 50–71.

Ashby, F. G., & Townsend, J. T. (1986). Varieties of perceptual independence. *Psychological Review, 95,* 154–179.

Buchner, A., & Wippich, W. (2000). On the reliability of implicit and explicit memory measures. *Cognitive Psychology, 40,* 227–259.

Busemeyer, J. R., Dewey, G. I., & Medin, D. L. (1984). Evaluation of exemplar-based generalization and the abstraction of categorical information. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 10,* 638–648.

Cleeremans, A., & McClelland, J. L. (1991). Learning the structure of event sequences. *Journal of Experimental Psychology: General, 120,* 235–253.

Elman, J. L. (1990). Finding structure in time. *Cognitive Science, 14,* 179–212.

Erickson, M. A., & Kruschke, J. K. (1998). Rules and exemplars in category learning. *Journal of Experimental Psychology: General, 127,* 107–140.

Estes, W. K. (1994). *Classification and cognition.* London: Oxford University Press.

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.

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

Hintzman, D. L. (1990). Human learning and memory: Connections and dissociations. *Annual Review of Psychology, 41,* 109–139.

Homa, D. (1984). On the nature of categories. In G. H. Bower (Ed.), *The psychology of learning and motivation* (Vol. 18, pp. 49–94). San Diego, CA: Academic Press.

Johansen, M. K., & Palmeri, T. J. (in press). Representational shifts in category learning. *Cognitive Psychology.*

Kinder, A., & Shanks, D. (2001). Amnesia and the declarative/nondeclarative distinction: A recurrent network model of classification, recognition, and repetition priming. *Journal of Cognitive Neuroscience, 13,* 1–22.

Knowlton, B. J., Mangels, J. A., & Squire, L. R. (1996). A neostriatal habit learning system in humans. *Science, 273,* 1399–1402.

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 categories: Parallel brain systems for item memory and category knowledge. *Science, 262,* 1747–1749.

Knowlton, B. J., & Squire, L. R. (1994). The information acquired during artificial grammar learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 20,* 79–91.

Knowlton, B. J., & Squire, L. R. (1996). Artificial grammar learning depends on implicit acquisition of both abstract and exemplar-specific information. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 22,* 169–181.

Kruschke, J. (1992). ALCOVE: An exemplar-based connectionist model of category learning. *Psychological Review, 99,* 22–44.

Love, B. C., Medin, D. L., & Gureckis, T. M. (in press). SUSTAIN: A network model of category learning. *Psychological Review.*

Maddox, W. T., & Ashby, F. G. (1993). Comparing decision bound and exemplar models of categorization. *Perception & Psychophysics, 53,* 49–70.

Medin, D. L. (1986). Commentary on "Memory storage and retrieval processes in category learning." *Journal of Experimental Psychology: General, 115,* 373–381.

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

Moscovitch, M. (1992). Memory and working-with-memory: A component process model based on modules and central systems. *Journal of Cognitive Neuroscience, 4,* 257–267.

Moscovitch, M. (1994). Memory and working with memory: Evaluation of a component process model and comparisons with other models. In D. L. Schacter & R. Tulving, Eds., *Memory systems 1994* (pp. 269–310), Cambridge, MA: MIT Press.

Myung, I. J. (2000). The importance of complexity in model selection. *Journal of Mathematical Psychology, 44,* 190–204.

Nosofsky, R. M. (1984). Choice, similarity, and the context theory of classification. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 10,* 104–114.

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

Nosofsky, R. M. (1988). Exemplar-based accounts of relations between classification, recognition, and typicality. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 14,* 700–708.

Nosofsky, R. M. (1991). Tests of an exemplar model for relating perceptual classification and recognition memory. *Journal of Experimental Psychology: Human Perception and Performance, 17,* 3–27.

Nosofsky, R. M., & Johansen, M. K. (2000). Exemplar-based accounts of "multiple-system" phenomena in perceptual categorization. *Psychonomic Bulletin & Review, 7,* 375–402.

Nosofsky, R. M., & Palmeri, T. J. (1997). An exemplar-based random walk model of speeded classification. *Psychological Review, 104,* 266–300.

Nosofsky, R. M., Palmeri, T. J., & McKinley, S. C. (1994). Rule-plus-exception model of classification learning. *Psychological Review, 101,* 53–79.

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. (1999). Math modeling, neuropsychology, and category learning: Response to B. Knowlton (1999). *Trends in Cognitive Sciences, 3,* 125–126.

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 under review.*

Palmeri, T. J. (1997). Exemplar similarity and the development of automaticity. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 23,* 324–354.

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.

Palmeri, T. J., & Nosofsky, R. M. (2001). Central tendencies, extreme points, and prototype enhancement effects in ill-defined perceptual categorization. *The Quarterly Journal of Experimental Psychology, 54,* 197–235.

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, A. S. (1969). Transfer of syntactic structure in synthetic languages. *Journal of Experimental Psychology, 81,* 115–119.

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

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.

Reber, P. J., Stark, C. E. L., & Squire, L. R. (1998a). Cortical areas supporting category learning identified using functional MRI. *Proceedings of the National Academy of Sciences USA, 95,* 747–750.

Reber, P. J., Stark, C. E. L., & Squire, L. R. (1998b). Contrasting cortical activity associated with category memory and recognition memory. *Learning and Memory, 5,* 420–428.

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.

Rickard, T. C., & Grafman, J. (1998). Losing their configural mind: Amnesic patients fail on transverse patterning. *Journal of Cognitive Neuroscience, 10,* 509–524.

Roediger, H. L., Buckner, R. L., & McDermott, K. B. (1999). Components of processing. In J. K. Foster & M. Jelicic (Eds.), *Memory: Systems, process, or function?* (pp. 31–65). Oxford, England: Oxford University Press.

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

Shin, H. J., & Nosofsky, R. M. (1992). Similarity-scaling studies of dot-pattern classification and recognition. *Journal of Experimental Psychology: General, 121,* 278–304.

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, and 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 Sciences USA, 92,* 12,470–12,474.

Squire, L. R., & Zola, S. M. (1996). Structure and function of declarative and nondeclarative memory systems. *Proceedings of the National Academy of Sciences USA, 93,* 13,515–13,522.

Vandierendonck, A. (1995). A parallel rule activation and rule synthesis model for generalization in category learning. *Psychonomic Bulletin & Review, 2,* 442–459.