Cognitive, Affective, & Behavioral Neuroscience
2001, 1 (4), 360-370

PET evidence for multiple strategies of categorization

ANDREA L. PATALANO, EDWARD E. SMITH, JOHN JONIDES, and ROBERT A. KOEPPE
University of Michigan, Ann Arbor, Michigan

It is widely held that people use multiple strategies to categorize their experiences in the world. We conducted a pair of neuroimaging experiments to identify the neural correlates of two of these strategies—rule application and exemplar similarity. Participants were instructed to perform either a rule- or an exemplar-based categorization task while changes in cerebral blood flow were measured using positron emission tomography. Patterns of neural activity were consistent with the predictions of cognitive models of rule- versus exemplar-based categorization and with existing neuroscience data. The identification of strategy-specific neural patterns offers future researchers a diagnostic tool for assessing strategy use in other situations.

Categorization is a fundamental human capacity that allows us to organize our knowledge and to use that knowledge to make inferences in novel situations. Two widely recognized categorization strategies[1] are rule application (e.g., Bruner, Goodnow, & Austin, 1956) and exemplar similarity (e.g., Estes, 1994; Nosofsky, 1992). Rule application involves assigning an entity to a category on the basis of a known rule that describes the category (e.g., "this is a hummingbird because it is red, eats sugar water, and can fly in place"). Exemplar similarity consists of assigning an entity to a category on the basis of the entity's similarity to previously encountered members of the category (e.g., "this must be a hummingbird because it looks like hummingbirds I have seen before"). Considerable behavioral evidence suggests that people have access to both of these strategies (e.g., Allen & Brooks, 1991; J. D. Smith & Kemler, 1984). The main goal of this paper is to identify the neural correlates of each strategy.

In addition to a better understanding of both cognitive and neural functioning, a second motivation for this work stems from the behavioral categorization literature. Although it is widely held that people have access to both rule application and exemplar similarity strategies (i.e., can use either when instructed), the extent to which each is used spontaneously is still the subject of much debate. At the extremes, a few categorization researchers have proposed models whereby much relevant behavioral data can be explained by appealing to either rule application (e.g., Nosofsky, Palmeri, & McKinley, 1994) or exemplar similarity (e.g., Estes, 1994). At this point, converging neuroscience evidence may be useful in addressing the conditions under which each strategy is spontaneously used. More specifically, if different strategies produce different patterns of neural activity, it may be possible to use neuroscience data in the future to help "diagnose" strategy use in controversial situations.

In recent years, psychologists have developed schematic models of the basic cognitive processes (e.g., working memory, selective attention, etc.) that may underlie rule application and exemplar similarity. At the same time, neuroscientists have identified some of the brain mechanisms that give rise to each of these basic processes. This means that it is possible to predict not only that use of rule application versus exemplar similarity strategies will lead to different patterns of brain activation, but more specifically, that these patterns should be consistent with hypothesized schematic cognitive models. According to existing cognitive models of visual categorization tasks,[2] rule application involves maintaining an explicit rule in working memory, selectively attending to each precondition of the rule, perceptually testing whether or not a target has the feature specified by the precondition, holding the results of each test in working memory, and combining the results to arrive at a categorization (e.g., Kemler-Nelson, 1984; E. E. Smith & Sloman, 1994). Visual selective attention is supported by the bilateral superior posterior parietal cortex (particularly Brodmann's area (BA) 7; e.g., Bushnell, Goldberg, & Robinson, 1981). The manipulation of information in working memory has been associated with the prefrontal cortex (including BAs 6, 10, 11, and 46), as well as with the superior posterior parietal cortex (e.g., Jonides et al., 1997). And basic visual perceptual processing has been associated with the occipital cortex bilaterally (BAs 17, 18, and 19; e.g., Kosslyn et al., 1993). Thus, one would

This research was funded by a grant from the Office of Naval Research. We thank Joe Bates and Chris Willis for their assistance in conducting this research. All authors are in the Department of Psychology, except Robert A. Koeppe, who is in the Department of Internal Medicine. Since conducting this research, Andrea L. Patalano has moved to the Beckman Institute at the University of Illinois, Champaign-Urbana. Correspondence concerning this article should be addressed to A. L. Patalano, Beckman Institute, University of Illinois, 405 N. Mathews Ave., Urbana, IL 61801 (e-mail: patalano@uiuc.edu).

Copyright 2001 Psychonomic Society, Inc.

expect to see activation in all of these areas during rule application categorization, although, of course, the degree of activity in each of these areas may depend on the particular rule (e.g., some are more working memory intensive than others).

Exemplar similarity, by contrast, involves visual processing of the target entity, retrieving from long-term memory stored instances that are similar to the target, and then assigning the target to the same category as the most similar instances (e.g., Estes, 1994; Medin & Schaffer, 1978; Nosofsky, 1986).[3] Both visual perceptual processing and long-term memory retrieval processes have been associated with the occipital cortex bilaterally (BAs 17, 18, and 19; e.g., Farah, McMullen, & Meyer, 1991; Kosslyn et al., 1993; Reber, Stark, & Squire, 1998), leading to a prediction of occipital activation during categorization by exemplar similarity. Of course, other cognitive processes may be minimally involved in either strategy (e.g., long-term memory retrieval of the rule during rule application or selectively placing greater weight on some stimulus features during similarity computations), but these should play a minor role, as compared with more central processes.

No known neuroscience studies specifically address the *use* of rule application or exemplar similarity in categorization. However, numerous patient studies (e.g., patients with frontal lobe damage; Milner, 1964) and functional magnetic resonance imaging (fMRI) studies (e.g., Elliott, Rees, & Dolan, 1999; Prabhakaran, Smith, Desmond, Glover, & Gabrieli, 1997; Rao et al., 1997) have investigated rule *learning* in reasoning contexts (e.g., with Raven's Progressive Matrices Task; Prabhakaran et al., 1997), whereas others have investigated non-rule-based learning in categorization contexts (e.g., Knowlton, Mangels, & Squire, 1996; Kolodny, 1994; Poldrack, Prabhakaran, Seger, & Gabrieli, 1999). Non-rule-based learning studies have broadly involved the gradual accumulation of information extracted from individual exemplars, but it is debated as to how many distinct strategies—such as exemplar similarity, cognitive skill learning, and perceptual priming—are captured by them. The relationship between the results of these studies and the present ones will be addressed in the General Discussion section.

Our methodology and behavioral analyses follow Allen and Brooks (1991), adapted to a positron emission tomography (PET) context. Their study offers a situation in which a simple instructional manipulation during training led to differential behavioral performance during a subsequent test phase consistent with the previously described schematic models of rule application and exemplar similarity (see E. E. Smith, Patalano, & Jonides, 1998, for a more extensive review of the literature). In a training phase, participants either were taught a complex rule (rule group) or were asked to use their memory of instances to learn to categorize two kinds of artificial animals (exemplar group), with feedback on each trial. In a subsequent test phase, designed to assess strategy use, participants were presented with similar but novel test animals. *Positive test-items* were those for which use of either strategy led to the same category assignment. *Negative test-items* were those for which use of each strategy led to a different category assignment. Category assignment and response time (RT) measures were collected.

Allen and Brooks (1991) found the following evidence for differential category use. First, the rule group categorized the majority of both kinds of test items according to the rule, whereas the exemplar group categorized the majority according to exemplar similarity. Second, the rule group responded more slowly overall, consistent with rule application's being a working-memory-intensive, multistage procedure. And third, the rule group was slower and made more errors (i.e., non-rule categorizations) in categorizing negative test items, as compared with positive test items, whereas the exemplar group data showed no difference between the two kinds of items. The latter is consistent with the idea that because rule application is a multistage, working-memory-intensive process, it is subject to greater response interference from other strategies, such as exemplar similarity. Overall, the data provide a sound argument that, at least in this context, distinct rule application and exemplar similarity strategies are being elicited.

Although a preliminary description of Experiment 1 appeared in E. E. Smith et al. (1998), that description contained minimal procedural details and only an initial (pixel-based) analysis of the PET data. Here, we provide complete procedural details, as well as hypothesis-driven analyses of the PET data.

EXPERIMENT 1

In our initial experiment, participants first learned to categorize a small set of artificial-animal stimuli, using either a given rule (rule group) or memorization of individual exemplars (exemplar group). The participants were then given novel test items to categorize and were told either to continue to rely on the rule (rule group) or to rely on previously studied animals (exemplar group). Throughout the study, category assignment and RT measures were collected. During test-item categorization, PET data were collected as well.

Method

Participants. Twenty-four 18- to 30-year-old University of Michigan undergraduates (12 male and 12 female) were recruited through newspaper advertisements and were paid for their participation. All were right-handed, native-English speakers with no neurological impairments; all consented to participate after a full briefing about the procedure. The participants were equally divided between rule and exemplar groups.

Materials. Artificial-animal stimuli were created (see Figure 1) that varied on 10 dimensions (e.g., body markings, tail shape) and adopted one of two features on each dimension (e.g., spots vs. stripes, straight vs. curly). Ten animals were designated training animals and were divided into two categories (with five animals per category) on the basis of the following rule: "An animal lives on Venus if it has at least three out of the following five features:

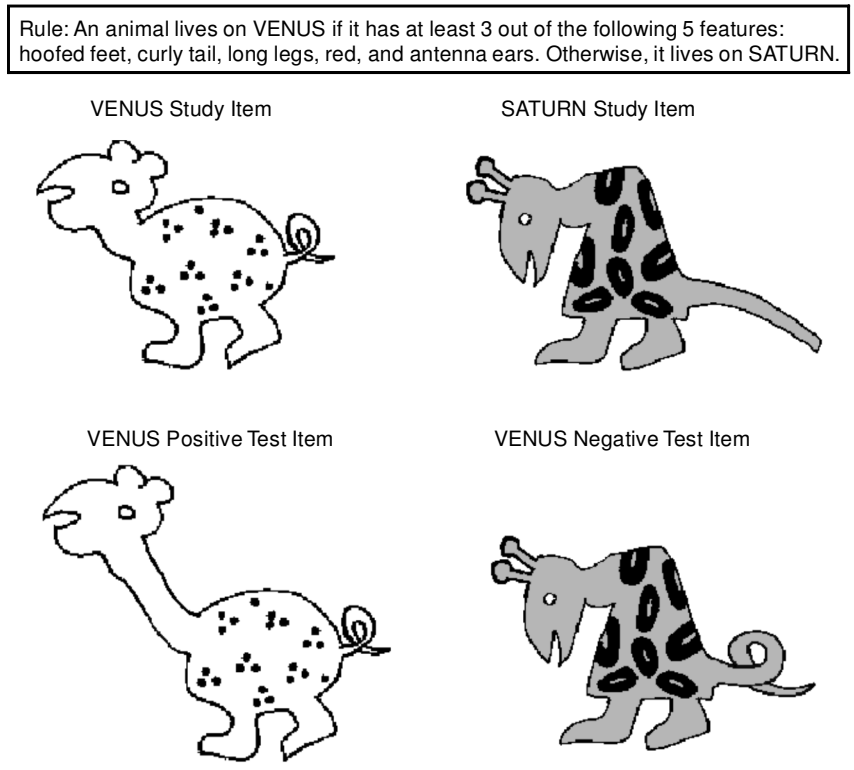


Figure 1. Example animal stimuli used in Experiments 1 and 2. In the original, the two left-hand images are red, and the two right-hand images are blue.

hoofed feet, a curly tail, long legs, red, antenna ears. Otherwise it lives on Saturn." The remaining animals were designated test animals. Each test animal was highly similar to one corresponding training animal in that it differed from the corresponding training animal by two features and differed from the other nine training animals by three or more features (see Table 1 for the structure of the stimulus set).

As in Allen and Brooks (1991), half of the test animals were *positive test items*: animals that should be assigned to the same categories regardless of whether one relied on rule application or exemplar similarity. The remaining test-items were *negative test items*: animals that fit one category according to the rule but were more similar to a training animal in the other category (e.g., an item that fits the rule for Venus but looks like a training item from Saturn). See Figure 1 for examples of positive and negative test animals.

General procedure. There were two sessions: a training and a PET session. During training, the 10 training animals were presented one at a time on a computer screen, the entire set being repeated eight times, each time in a different random order. On each trial, the participant categorized the animal, studied feedback regarding the animal's actual category, and pressed a button to go on to the next animal. The first four sets of animals were presented with an unlimited categorization time (the animal remained on the screen until a response was made). The second four sets were presented with a 2.5-sec exposure time and a 0.5-sec interstimulus interval; only responses made during this 3-sec time window were recorded, although feedback study time remained unlimited. The rule and the exemplar conditions were identical, except in their instructions. Rule instructions provided the rule and told the participants to use it to categorize all training items. Exemplar instructions did not present the rule and told the participants to "use your memory of previously studied animals to help you."

The training session also included 20 practice test trials. Test trials were identical to the timed training trials, except that feedback was omitted and training stimuli were replaced with novel test stimuli. The participants were instructed to continue to use the same strategy (rule application vs. exemplar similarity) to categorize the animals. Finally, the participants filled out a questionnaire in which they were asked to describe the strategy they had used to categorize the animals. All 24 participants met the criteria of (1) at least 60% accuracy on practice test trials (accuracy was assessed relative to assigned condition) and (2) condition-appropriate questionnaire responses (i.e., rule participants reported having used a rule application strategy, and exemplar participants an exemplar similarity strategy).[4] All the participants returned for the PET session within 4 days.

During the PET session, the participants completed one practice timed-training block, followed by three test blocks and three control blocks (all blocks of 20 items; blocks were counterbalanced), while being scanned. The control blocks were identical to the test blocks, except that the participants were instructed not to categorize the animals but just to look at each animal and press either response button.

PET injection procedure. Before beginning each scanned block, the participant was injected with 40 m Ci of ^{15}O water. The participant then started the task and completed approximately 8 trials while waiting for the ^{15}O water to reach the brain. Collection of PET data began 5 sec after the activity in the brain rose above the background level and continued for 60 sec. For the first 35 sec (until the approximate point of peak activity in the brain), the participant continued to respond to animal stimuli, completing 12 trials. Injections occurred 12 min apart—the point at which clearance from the brain plus radioactive decay decreased to less than 1% of background level. Data were collected using a Siemens EXACT-47 PET scanner.

Table 1
Stimulus Materials for Experiments 1 and 2

Category A										Category B									
NL	TS	FS	FF	ES	CL	BS	BM	HD	LL	NL	TS	FS	FF	ES	CL	BS	BM	HD	LL
Training Items (All Items)																			
1	2	2	1	1	1	1	1	1	1	2	1	2	1	2	2	2	1	1	1
2	2	1	1	1	2	2	2	2	2	1	2	2	2	1	2	1	1	1	2
1	1	1	2	2	2	2	1	1	2	2	1	1	2	2	1	1	2	2	2
1	1	2	2	1	1	2	2	2	1	1	2	2	1	2	1	2	2	1	2
2	1	1	1	2	1	1	1	1	2	2	2	1	2	1	1	2	1	2	1
Positive Test Items (Sample Only)																			
1	2	2	1	1	1	2	1	1	1	2	1	2	1	2	1	2	1	1	1
2	2	1	1	1	2	2	1	2	2	1	2	2	2	1	2	2	1	1	2
1	1	1	2	2	1	2	1	1	2	2	1	1	2	2	1	1	2	1	2
1	1	2	2	1	2	2	2	2	1	1	2	2	1	2	1	2	2	2	2
2	1	1	1	2	1	1	1	2	2	2	2	1	2	1	2	2	1	2	1
Negative Test Items (Sample Only)																			
2	2	2	1	1	1	1	1	1	1	2	1	2	1	1	2	2	1	1	1
2	2	1	2	1	2	2	2	2	2	1	2	1	2	1	2	1	1	1	2
2	1	1	2	2	2	2	1	1	2	1	1	1	2	2	1	1	2	2	2
1	1	2	2	2	1	2	2	2	1	1	2	2	1	1	1	2	2	1	2
2	1	2	1	2	1	1	1	1	2	2	1	1	2	1	1	2	1	2	1

Note—Stimulus dimensions and values are, in the order presented above, neck length (NL; 1 = long, 2 = short), tail shape (TS; curly, straight), foot shape (FS; hoofed, webbed), facial feature (FF; beak, snout), ear shape (ES; antennae, rounded ears), color (CL; red, blue), body shape (BS; rounded, angular), body markings (BM; stripes, spots), head direction (HD; up, down), and leg length (LL; long, short). The first five dimensions are relevant to the rule (items with at least 60% 1s belong in Category A according to the rule), whereas the second five are irrelevant. Each test stimulus is listed under the category to which it should be assigned on the basis of exemplar similarity.

Results

Behavioral results. Because the patterns of behavioral results for training and test trials are highly similar, results are reported for scanned test trials only. Following Allen and Brooks (1991), three indicators of differential strategy use were considered. First, the rule group categorized predominantly according to rule application, whereas exemplar group categorized according to exemplar similarity (78% vs. 77%, respectively). Second, rule group RTs were slower than exemplar group RTs [1,907 vs. 1,144 msec; $t(22) = 7.71$, $p < .001$]. And third, a group × trial type (positive vs. negative) interaction was significant for accuracy levels (but not for response times, $p > .1$, unlike in Allen & Brooks, 1991). Rule group accuracy was 85% versus 71% (positive vs. negative items), whereas exemplar group accuracy was 77% versus 76%, respectively [for the interaction, $F(1,23) = 6.75$, $p = .016$]. Recall that the positive–negative distinction is meaningless for the exemplar group—it is simply a division of stimuli according to whether they are positive or negative for the rule group—so one would not expect to see any effect of item type for the exemplar group. The rule-group accuracy results may seem low, especially for a task as straightforward as applying a given rule, but the strict time limit (of 3 sec) increased task difficulty. Overall, the results replicate those of Allen and Brooks and confirm that the participants were using the instructed strategies.

PET results: General findings. The scanner produced 47 contiguous brain slices 3.375 mm center-to-center apart, with a spatial resolution of 10 mm for each scan. To create subtraction images, intraparticipant registration by an automatic algorithm corrected for participant motion across scans. Then, individual-participant image sets were transformed to a standard bicommissural stereotactic system through detection of mid-sagittal plane, detection of bicommissural (AC–PC) line of the mid-sagittal plane, and linear scaling followed by nonlinear warping of the brain to remove anatomic differences among participants (Minoshima, Berger, Lee, & Mintun, 1992; Minoshima et al., 1993). And finally, a subtraction image set was created for each participant by using the average of the categorization scans minus the control scans. The subtraction image sets, then averaged across participants, produced a group average subtraction set consisting of a mean and a standard deviation of cerebral blood flow (CBF) for each brain voxel.[5] Standard deviations for the voxels were averaged within the brain to create a pooled estimate of variance, and *t*-statistical values were computed for each voxel using a pooled variance estimate and adjusting for multiple nonindependent comparisons (Friston, Frith, Liddle, & Frackowiak, 1991; Worsley, Evans, Marrett, & Neelin, 1992). A one-tailed adjusted value of $p < .05$ is used as criterion of reliability for peak-voxel analyses throughout the paper.

The peak voxels in each active region were identified. Twenty-one peak voxels were identified for the rule group, and nine were identified for the exemplar group. These peaks are listed in Table 1 by stereotactic coordinate and by associated BA. Rule versus exemplar peaks are aligned in the table if they are within 11 mm of one

Table 2
Rule Versus Exemplar Conditions in Experiment 1:
Voxels of Peak Activation Identified by Brodmann's Area (BA), Stereotactic Coordinates, and *Z* Score

Rule Condition					Exemplar Condition				
Region	*x*	*y*	*z*	*Z*	Region	*x*	*y*	*z*	*Z*
BA 46 R	−39	19	25	5.38					
BA 6 L	30	−4	47	5.01					
BA 6 L	37	−6	45	4.95					
BA 6 R	−26	−6	45	5.31					
BA 7/19 L	24	−60	40	8.40					
BA 7/19 R	-19	−64	40	7.82					
BA 19 L	28	−80	16	7.01					
BA 19 L	42	−67	−9	5.47					
BA 19 R	−35	−73	20	5.84					
BA 19 R	−35	−60	−9	5.80	BA 19 R	−26	−67	−11	6.50
BA 18 L	1	−69	2	7.77	BA 31 L	3	−71	7	6.36
BA 18 L	26	−82	−4	7.35	BA 18 L	26	−78	−7	7.72
BA 18 R	−26	−82	2	6.24					
BA 18 R	−12	−76	−7	6.98	BA 18 R	−6	−82	−9	7.13
BA 17 L	15	−85	7	7.49	BA 17 L	12	−89	4	6.67
Cb L	10	−64	−16	6.19					
Cb R	−8	−62	−14	6.99					
Cb R	−3	−55	−22	7.31	Cb R	−17	−51	−14	4.83
Cb R	−6	−69	−25	7.53	Cb R	−1	−67	−27	6.85
Cb R	−26	−40	−32	4.75					
Th R	−1	−13	0	4.72					
					BA 18 L	10	−82	−11	6.94
					Cb L	37	−46	−16	4.46

Note—Cb, cerebellum; Th, thalamus. Voxels are matched across columns if they are within 11 mm on each coordinate.

another on each coordinate (approximately the spatial resolution of our PET data). The two groups showed corresponding peaks in the occipital cortex in BAs L17 (area 17 of the left hemisphere), L18, R18, R19, and RCb[2] (two different peaks in the right cerebellum). In the rule condition, additional peak activations were found in the superior posterior parietal cortex in BAs L7/19 (at the junction of the two areas) and R7/19, and in the prefrontal cortex in R6, L6, and R46. A number of other peaks in areas 18, 19, the cerebellum, and the thalamus were active in one condition or the other (see Table 2 for details).

To compare the activation patterns for the two groups, the following *region-of-interest* (ROI) analysis was used. First, the rule and the exemplar subtraction images were combined and analyzed (across all participants) for significant voxels of activation. For each of the 18 peak voxels identified—of which 8 were in occipital cortex (L17, L18, R18[2], L19[2], R19[2]), 2 were in the superior posterior parietal cortex (L7/19 and R7/19), 4 were in the prefrontal cortex (R46, R10[2], and R11), and 4 were in other areas (L37, RCb[2], and RNr [right nucleus ruber])—a spherical ROI with an 11-mm radius was defined around the voxel. The average activation within each ROI was then compared for rule versus exemplar groups. Significantly greater activation was found in the rule than in the exemplar group (using a one-tailed uncorrected $p < .05$ criterion for ROI analyses throughout this paper) in the left superior posterior parietal cortex [L7/19; stereotactic coordinates, $x = 24$, $y = -62$, $z = 40$; $t(23) = 3.38$, $p = .001$] and the right superior parietal cortex [R7/19; $x = -21$, $y = -64$, $z = 38$; $t(23) = 1.98$, $p = .030$]. The right prefrontal cortex [R46; $x = -48$, $y = 28$, $z = 25$; $t(23) = 1.61$, $p = .060$] was also nearly statistically significant.

Are these three ROI differences qualitative or quantitative only? That is, do they reflect increased activity in the rule group, as compared with *no* increase in activity in the exemplar group, or do they simply reflect *less* increase in the exemplar group? To address this question, we considered whether or not each of the three relevant ROI activations applied to the exemplar group was significantly different from 0% change in CBF. In the prefrontal cortex, exemplar group activation was not different from 0% [$M = 3.4\%$; $t(11) = 1.49$, $p = .100$], suggesting a qualitative difference between the rule and the exemplar groups. In both left ($M = 3.4\%$) and right ($M = 5.6\%$) posterior parietal areas, respectively, exemplar activation was greater than 0% ($p < .050$), as it was in all three rule ROIs (prefrontal $M = 6.5\%$; left parietal $M = 10.8\%$; right parietal $M = 9.8\%$; $p < .050$). Although the latter is consistent with a quantitative difference only, it is a strikingly large difference: 7.4% change in CBF for the left parietal area and 4.2% for the right parietal area.

Overall, these neuroimaging results are consistent with the previously elaborated models of rule application and exemplar similarity and with existing neuroscience literature. Furthermore, they suggest two distinct neural circuits underlying rule application and exemplar similarity, respectively. The latter is important both because it fits with previous models and literature and because it suggests that these patterns might be used to diag-

nose strategy use in future work. Despite the consistency of these results, some further analyses were conducted to address plausible alternative explanations.

PET results: Time-on-task analyses. An alternative interpretation is that the increased activation in the working memory and selective attention regions for the rule group is simply due to the increased time required to perform the rule task (an additional 650 msec on each trial), as compared with the exemplar task. In other words, the two tasks may activate the same areas, but because the rule task is more time intensive, it activates these areas for a longer time over the course of a scan. To test this, we created four subgroups—slow-rule (mean RT = 2,196 msec), fast-rule (1,488 msec), slow-exemplar (1,468 msec), and fast-exemplar (802 msec)—consisting of the two fastest and slowest participants in each condition. Subgroups of this size were chosen to equate fast-rule and slow-exemplar subgroup response times and to ensure large RT differences between fast and slow subgroups within a condition. For each subgroup, we computed the average activation across all working memory and selective attention ROIs (all prefrontal and parietal areas). According to the time-on-task explanation, there should be no difference in ROIs for RT-matched fast-rule and slow-exemplar subgroups, and slow subgroups in both conditions should show greater average activation than do fast subgroups. The data contradict both of these predictions. The fast-rule subgroup showed significantly greater activation than did the slow-exemplar subgroup [6.2% vs. 2.4%; $t(3) = 3.17$, $p = .050$], the slow- and fast-rule subgroups did not differ in activation [6.1% vs. 6.2%; $t(3) = 0.03$, $p = .980$], and the slow-exemplar subgroup had *less* activation than did the fast subgroup [2.4% vs. 6.9%; $t(3) = 7.00$, $p = .006$]. Although we do not have an explanation for the anomalous fast-exemplar subgroup mean, the results argue against a time-on-task explanation, as do the results of Experiment 2.

PET results: Control task analyses. The final analysis addressed an issue of potential concern regarding the initially described voxel-by-voxel subtractions themselves. Given that exemplar similarity is hypothesized to require fewer cognitive stages and to be less working-memory-intensive than is rule application, it is possible that the exemplar group—but not the rule group—engaged in categorization during the control task, despite the instructions. If this were the case, some neural activity of interest would have been subtracted from the exemplar experimental condition. To address this possibility, we directly compared the raw experimental condition activations (i.e., without subtraction of associated control conditions), using an unpaired, one-tailed *t* test statistic, and did the same for the two raw control conditions. Evidence for this alternative interpretation of our findings would consist of a significant difference between the control conditions and no difference between the experimental conditions. Evidence for our original interpretation would be the reverse—namely, a significant difference between the experimental conditions only.

To counter the increased statistical noise associated with a between-subjects comparison (relative to the standard within-subjects comparisons), the data were first transformed by centering spherical ROIs of 2-cm diameter on the voxels identified as significantly activated in the original within-condition experimental minus control subtractions (shown in Table 2). These ROIs were then grouped into six distinct brain regions—superior parietal cortex, dorsolateral prefrontal cortex, medial prefrontal cortex, visual cortex, cerebellum, and thalamus (the first three of which were active only in the rule group in the original analysis, the final three being active in both groups). All ROIs within each region were averaged together. A comparison between the rule and the exemplar experimental conditions revealed that the rule condition was significantly more active in the three expected areas: superior parietal cortex [$t(22) = 2.23$, $p = .018$], dorsolateral prefrontal cortex [$t(22) = 1.97$, $p = .031$], and medial prefrontal cortex [$t(22) = 1.76$, $p = .046$]. No other differences were found in either the experimental or the control group comparison [$t < 0.7$, $p > .25$]. The results are inconsistent with the argument that differences in the original subtractions can be attributed in any way to differences in the control conditions.

Discussion

The behavioral results replicate past studies in that differential RTs, responses to negative test items, and responses to questionnaires provide evidence that participants in the rule and exemplar groups were engaged in distinct categorization procedures. More strikingly, the neuroimaging data revealed distinct neural patterns for the two strategies. Rule application activated the prefrontal (BAs 6 and 46), the posterior parietal (BA 7), the occipital (BAs 17, 18, and 19), and the subcortical areas (thalamus and cerebellum). Exemplar similarity activated the occipital (BAs 17, 18, and 19) and the subcortical (cerebellum) areas only. In other words, a selective dissociation was found between the rule and the exemplar groups, most notably owing to activation in the rule condition in areas previously associated with working memory and selective attention.

In Experiment 1, the rule and the exemplar conditions were made as similar as possible so that any differences in activation could be attributed only to differences in categorization strategy. As a result, the rule participants may have memorized the training items and engaged in some exemplar-based categorization of test items. This is suggested by the behavioral finding first identified by Allen and Brooks (1991) and partially replicated here, that rule group performance can be slower and less accurate for negative than for positive test items, consistent with the notion of interference from exemplar similarity processes during rule application. In Experiment 1, interference is likely to have only minimally "contaminated" the neuroimaging results by possibly inflating the amount of occipital activation during rule application (the region active in the exemplar condition). Nonetheless, to isolate

the brain areas involved in rule application only, we performed a second PET experiment, in which exemplar similarity was no longer an available strategy during rule application. To assess the contribution of exemplar similarity processes to rule group occipital activation in Experiment 1, we looked for a decrease in occipital activation from Experiment 1 to Experiment 2, relative to other brain regions. The present study also afforded a second opportunity to assess the time-on-task explanation of between-conditions differences in activation.

EXPERIMENT 2

Experiment 2 replicated the rule condition of Experiment 1, except that one change was made to decrease any use of an exemplar similarity strategy. For each test block, the participants were presented with a novel categorization rule. As a result, the participants could not come to associate test animals with particular categories (because an animal's category assignment was continually changing), and thus, participants would be unlikely to retrieve old items when categorizing new items. Behaviorally, this should result in a decrease in the accuracy difference between positive and negative items, and evidence of such a decrease could serve as a manipulation check. Overall, we expected the results of Experiment 2 to replicate those of Experiment 1. However, if the rule results of Experiment 1 were contaminated by exemplar similarity processes, Experiment 2 should show decreased activation in the occipital cortex (i.e., the area active in the exemplar condition of Experiment 1), relative to other regions (e.g., those involved in selective attention and working memory).

Method

Participants. Nine 18- to 30-year-old University of Michigan undergraduates (4 male and 5 female) were recruited through newspaper advertisements and were paid for their participation. All were right-handed and native English speakers with no neurological impairments, and all consented to participate after a description of the procedure.

Materials and Procedure. The materials and procedure of Experiment 2 were the same as those in Experiment 1, with the following exceptions. First, additional test stimuli (using the same criteria as before) and test rules (with the same "at least three out of five" format) were created. Second, stimulus presentation time was increased from 2.5 to 3.5 sec. Because of this timing change, only 9 stimuli (as compared with 12 in Experiment 1) could be presented during the first 35 sec of scanning time. Third, during the training session on the 1st day of the study, after completing a sequence of trials like that in Experiment 1, the participants were given a second set of 20 practice test trials, but with a new rule, and they were instructed to use the new rule to categorize items during these test trials. All the participants scored at least 60% by using this new rule and indicated a rule-consistent strategy on the questionnaire. Fourth, on the 2nd day of the experiment (when the scans occurred), the participants were presented with a new rule for each block of categorization trials. Fifth, the participants completed four scanned test blocks and four scanned control blocks (rather than the three of each used in Experiment 1). And finally, injections followed a 10-min, rather than a 12-min, cycle, with 48 m Ci ^{15}O water injected per scan.

Results

Behavioral results. RTs averaged 2,683 msec during the scanned test trials (as compared with 1,907 msec in the rule condition of Experiment 1). Accuracy (i.e., category assignments based on rule application) averaged 65% (as compared with 78% in the rule condition of Experiment 1). Accuracy was presumably lower here because the rule kept changing, but all the participants still performed significantly above chance. Importantly, unlike Experiment 1, accuracy rates were the same for scanned positive versus negative test trials—66% versus 64%, respectively ($p > .1$)—and there continued to be no difference in RTs for these items (2,644 for positive items vs. 2,716 for negative items; $p > .1$). Thus, as was predicted, there was no evidence of interference from exemplar similarity processes in the present rule group.

PET results: General findings. PET data were analyzed as in Experiment 1. Schematics of the subtraction images are presented in Figure 2, overlaid on the rule condition results from Experiment 1. Eleven peak areas of activation were found in BAs L44/46, L6, L7/19, R7/19, R19, L18, R18, LCb, and RCb[3]. Region R9/46 was also marginally significant. These voxels are presented in Table 2, aligned with corresponding voxels from the rule group of Experiment 1. Note that although there are fewer peaks in Experiment 2, the overall pattern is strikingly similar to that of Experiment 1. Twenty-one ROIs were developed around the peak voxels from the rule condition of Experiment 1 (using the same technique as that described in Experiment 1) and placed on the Experiment 2 data; 16 of the 21 ROIs were again significant in Experiment 2 (see Table 3). Reciprocally, 12 ROIs were developed around the peak voxels from the present rule condition and were applied to the rule condition of Experiment 1; 11 of the 12 were significantly greater than 0 (see Table 3). Overall, the PET results replicate those of Experiment 1.

To test whether or not activation in occipital areas decreased relative to other areas, the same previously mentioned 21 ROIs developed around the Experiment 1 rule group data were applied to both Experiment 1 and Experiment 2. The average activations of selective attention and working memory areas (L7/19, R7/19, R46, R6, and L6[2]) versus visual areas (L17, L18[2], R18[2], L19[2], and R19[2]) were computed for each participant. Overall, the mean activations in Experiment 1—8.3% (working-memory/selective-attention areas) versus 8.2% (visual areas)—were greater than those in Experiment 2 [5.3% and 3.6%, respectively; $F(1,20) = 19.18, p < .001$]. The interaction was not statistically significant [$F(1,20) = 0.72, p = .404$]. The same main-effect-only pattern was found when the reciprocal analysis was performed by applying the Experiment 2 ROIs to both experiments. These results suggest that although memory retrieval processes may have led to a small increase, at most, in occipital activation in the rule condition of Experiment 1, it did not fundamentally alter the pattern of neural activation.

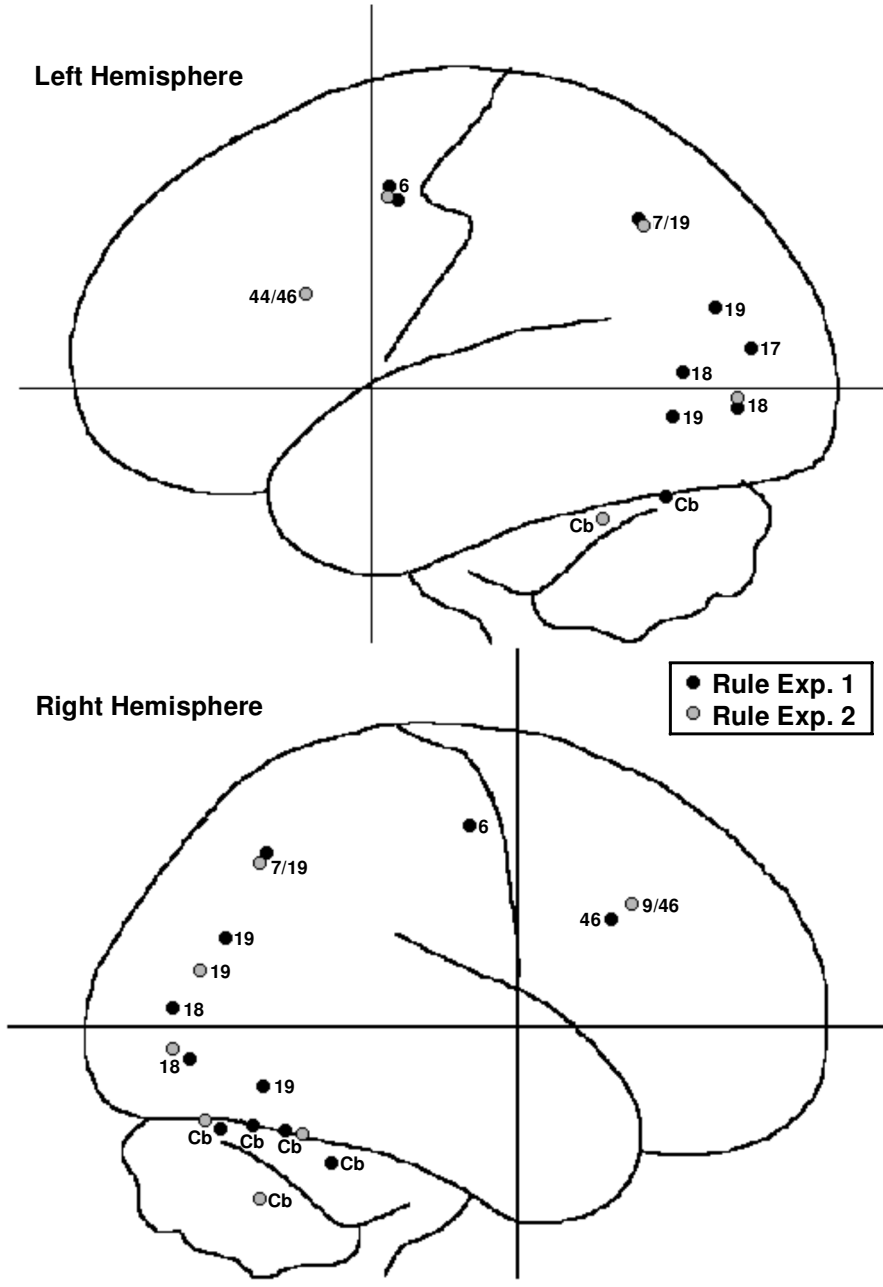


Figure 2. Schematic of Experiment 1 versus Experiment 2 rule condition voxels of peak activation.

Time-on-task analyses. As in Experiment 1, an additional comparison was made of the fastest versus slowest performers to test the hypothesis that increased activation was due to time spent on task. The ROIs used in Experiment 1 (from the combined rule and exemplar conditions) were also applied to Experiment 2, and an average activation level was again obtained for working memory and selective attention ROIs for the fastest (mean RT = 2,266 msec) vs. slowest (mean RT = 2,982 msec) 2 participants. In this experiment, the fast-rule group had *higher* levels of activation (5.4%) than did the slow-rule group (1.7%), although the difference was not statistically significant [$t(3) = 1.05, p = .370$].

Discussion

The present rule-condition neuroimaging results—obtained in a situation in which exemplar similarity was not an available strategy—replicate those of the rule group of Experiment 1. The results provide clear evidence of a neural circuit involved in rule application and suggest only marginal influence of exemplar similarity processes on the original rule condition results. The small change

Table 3
Rule Conditions in Experiment 1 Versus Experiment 2:
Voxels of Peak Activation Identified by Brodmann's Area (BA), Stereotactic Coordinates, and *Z* Score

Experiment 1 Rule Condition					Experiment 2 Rule 2 Condition				
Region	*x*	*y*	*z*	*Z*	Region	*x*	*y*	*z*	*Z*
BA 46 R	−39	19	25	5.38	BA 44/46 L	44	12	25	5.35*
BA 9/46 R	−55	28	29	4.10					
+BA 6 L	30	−4	47	5.01	BA 6 L	28	−4	45	4.80
+BA 6 L	37	−6	45	4.95					
BA 6 R	−26	−6	45	5.31					
+BA 7/19 L	24	−60	40	8.40	BA 7/19 L	28	−64	38	7.75
+BA 7/19 R	−19	−64	40	7.82	BA 7/19 R	−24	−64	38	7.65
+BA 19 L	28	−80	16	7.01					
+BA 19 L	42	−67	−9	5.47					
+BA 19 R	−35	−73	20	5.84	BA 19 R	−37	−80	9	5.56
+BA 19 R	−35	−60	−9	5.80					
BA 18 L	1	−69	2	7.77					
+BA 18 L	26	−82	−4	7.35	BA 18 L	26	−82	−2	8.68
+BA 18 R	−26	−82	2	6.24	BA 18 R	−26	−82	−4	7.53
+BA 18 R	−12	−76	−7	6.98					
BA 17 L	15	−85	7	7.49					
+Cb L	10	−64	−16	6.19					
					BA Cb L	35	−46	−16	4.57
+Cb R	−8	−62	−14	6.99	BA Cb R	−1	−71	−22	7.98
+Cb R	−3	−55	−22	7.31					
+Cb R	−6	−69	−25	7.53					
Cb R	−26	−40	−32	4.75					
+Th R	−1	−13	0	4.72					
					BA Cb R	−30	−51	−14	6.26
					BA Cb R	−33	−60	−40	5.42

Note—Cb, cerebellum; Th, thalamus. Voxels are matched across columns if they are within 11 mm on each coordinate. *Differs in hemisphere from matched Experiment 1 voxel. +Region of interest developed around voxel is significant in Experiment 2.

in neuroimaging results across the two experiments may seem odd, given the moderate change in behavioral results, but it is quite possible that different scales (i.e., accuracy or RT, on the one hand, and neural activation, on the other) map onto one another in complex ways. The results support the conclusion of Experiment 1 regarding time on task: There is no evidence that activation of frontal and parietal brain areas is a function of time spent on the categorization task during scanning.

GENERAL DISCUSSION

In these studies, categorization tasks involving rule activation and exemplar similarity activated two different neural circuits. These predicted neural patterns are consistent with schematic models of rule application and exemplar similarity strategies described in this paper and throughout the categorization literature (see E. E. Smith et al., 1998). Specifically, rule application tasks led to activation of the occipital cortex (consistent with the role of visual perceptual testing in rule application), the posterior parietal cortex (visual selective attention), and the prefrontal cortex (working memory). Exemplar similarity tasks activated areas predominantly in the occipital cortex (consistent with the role of perceptual processing and visual long-term memory retrieval in exemplar similarity). Of course, because the same occipital areas—at least given the resolution of PET—were activated in both experimental conditions, the condition-specific interpretation of the activations require further investigation.

The results of the rule application condition are consistent with previous neuroscience research on rule learning. Rule-learning tasks generally involve rule application processes plus additional learning-related processes (e.g., rule generation, selection, and switching). Not surprisingly, fMRI studies involving these tasks typically result in activation of the same circuit as that found here, plus additional areas presumably involved in learning (e.g., Elliott et al., 1999; Prabhakaran et al., 1997; Rao et al., 1997). For example, an fMRI-adapted version of Raven's Progressive Matrices Task activated the prefrontal, posterior parietal, and occipital cortices, plus the anterior cingulate (Prabhakaran et al., 1997). An adapted version of the WCST activated the same regions, plus numerous subcortical structures (including the thalamus and basal ganglia; Rao et al., 1997). It has been suggested that the anterior cingulate may play a role in rule selection, whereas the thalamus and basal ganglia may mediate rule switching (Ashby, Alfonso-Reese, Turken, & Waldron, 1998). The results of the present experiments provide further evidence for a circuit specific to rule application and support the distinctions that have been made between these areas and those involved in learning-related processes.

The relationship between our exemplar condition results and neuroscience research involving the accumula-

tion of information from exemplars is less clear. One view in the neuroscience of categorization is that exemplar learning is a procedural skill (e.g., Ashby et al., 1998). This claim is based on three major findings. First, anterograde amnesics, after being exposed to a series of variations on a category prototype (e.g., artificial animals), are able to appropriately categorize new variations despite severely impaired recognition memory for the original items (e.g., Reed, Squire, Patalano, Smith, & Jonides, 1999). Second, Parkinson's patients, who have difficulty learning procedural skills, are impaired at prototype-based and probabilistic category-learning tasks (Knowlton et al., 1996). And third, performance on similar tasks during fMRI produces striatal activation, which is also associated with motor skill learning (and is damaged in Parkinson's disease; Poldrack et al., 1999).

Although these arguments strongly suggest that some information about category exemplars is extracted using a procedurally based system, it is unlikely that this is the only system able to utilize exemplar information. Pickering (1997), for example, argues that element-based representations (i.e., associations of varying strength between individual features and categories) are encoded by the striatum but that the binding together of the features that make up an exemplar (needed for use of an exemplar similarity strategy) occurs in the hippocampus. It is our belief that the latter system is likely to be used when exemplars are complex, when they are distinctive (in relation to one another), and when there is a high similarity between each target item and a specific stored exemplar. Other than the exemplar condition of this paper, only one known study—in which amnesics categorized artwork by various fictitious painters—fits this description (Kolodny, 1994). The fact that amnesics performed at chance on this task and our failure to find striatal activation in our exemplar task are both evidence for a distinction between the two exemplar-learning systems.

It may be tempting to ask why medial temporal activation was not found in our exemplar condition. Although there is strong evidence that the medial temporal lobe is involved in exemplar learning (e.g., Reed et al., 1999), its activation in retrieval contexts has not been observed consistently (e.g., Nyberg et al., 1995; Reber et al., 1998; Rugg, Fletcher, Frith, Frambach, & Dolan, 1996). For example, no medial temporal activation was found in an fMRI task in which participants gave recognition judgements after studying five dot patterns eight times each (Reber et al., 1998), even though medial temporal damage impairs the learning of these stimuli (e.g., Kolodny, 1994). Although the reason for these inconsistencies is unclear, it is not surprising that medial temporal activation did not emerge here.

Ashby and colleagues (Ashby et al., 1998) have proposed a neuropsychological theory in which category learning is a competition between separate verbal-explicit and (nonverbal) implicit systems. According to the theory, the rule system involves the anterior cingulate for selecting among rules, the prefrontal cortex for applying rules, and the basal ganglia for switching among rules. Our experimental tasks involved the application of a single rule and resulted in prefrontal and posterior parietal activation and, thus, is consistent with Ashby et al.'s model, with the exception that their model does not account for the posterior parietal activation. The implicit system in the model involves the occipital cortex for creating high-level visual representations of stimuli, the striatum for learning associations between representations and responses, and the prefrontal cortex for maintaining response possibilities. We did not observe striatal or prefrontal activation in our exemplar similarity condition. It is possible that these areas may play less of a role in skill (or exemplar) *use* than in skill *learning* (Ashby et al., 1998, suggest this about the striatum in particular), but further research is needed to test this possibility.

The results of the present experiments do not directly contribute to our understanding of spontaneous strategy use (e.g., J. D. Smith & Kemler, 1984). Also, they are not informative about the nature of the strategies in different experimental contexts (e.g., Nosofsky et al., 1994) or strategy use in more complex categorization tasks, such as medical diagnosis (Brooks, Norman, & Allen, 1991). Thus, an important next step is to use the neural patterns identified in this paper as a tool for assessing strategy use in a broader range of categorization situations of interest.

REFERENCES

Allen, S. W., & Brooks, L. R. (1991). Specializing the operation of an explicit rule. *Journal of Experimental Psychology: General*, **120**, 3-19.

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

Brooks, L. R., Norman, G. R., & Allen, S. W. (1991). Role of specific similarity in a medical diagnostic task. *Journal of Experimental Psychology: General*, **120**, 278-287.

Bruner, J. S., Goodnow, J., & Austin, G. A. (1956). *A study of thinking.* New York: Wiley.

Bushnell, M. C., Goldberg, M. E., & Robinson, D. L. (1981). Behavioral evidence of visual response in monkey cerebral cortex: I. Modulation in posterior parietal cortex related to selective visual attention. *Journal of Neurophysiology*, **46**, 755-772.

Elliott, R., Rees, G., & Dolan, R. J. (1999). Ventromedial prefrontal cortex mediates guessing. *Neuropsychologia*, **37**, 403-411.

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

Farah, M., McMullen, P. A., & Meyer, M. M. (1991). Can recognition of living things be selectively impaired? *Neuropsychologia*, **29**, 185-193.

Friston, K. J., Frith, C. D., Liddle, P. F., & Frackowiak, R. S. J. (1991). Comparing functional (PET) images: The assessment of significant change. *Journal of Cerebral Blood Flow & Metabolism*, **11**, 690-699.

Jonides, J., Schumacher, E., Smith, E. E., Lauber, E., Awh, E., Minoshima, S., & Koeppe, R. A. (1997). Verbal working memory load affects regional brain activity as measured by PET. *Journal of Cognitive Neuroscience*, **9**, 467-475.

Kemler-Nelson, D. (1984). The effect of intention on what concepts are acquired. *Journal of Verbal Learning & Verbal Behavior*, **23**, 734-759.

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

Kolodny, J. A. (1994). Memory processes in classification learning:

An investigation of amnesic performance in the categorization of dot patterns and artistic styles. *Psychological Science*, **5**, 164-169.

Kosslyn, S. M., Alpert, N. M., Thompson, W. L., Malijkovic, V., Weise, S. B., Chabris, C. F., Hamilton, S. E., Raunch, S., Buonanno, F. S. (1993). Visual mental imagery activates topographically organized visual cortex: PET investigations. *Journal of Cognitive Neuroscience*, **5**, 263-287.

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

Milner, B. (1964). Some effects of frontal lobotomy on man. In J. M. Warren & K. Akert (Eds.), *The frontal granular cortex and behavior*. New York: McGraw-Hill.

Minoshima, S., Berger, K. L., Lee, K. S., & Mintun, M. A. (1992). An automated method for rotational correction and centering of three-dimensional functional brain images. *Journal of Nuclear Medicine*, **33**, 1579-1585.

Minoshima, S., Koeppe, R. A., Mintun, M. A., Berger, K. L., Taylor, S. F., Frey, K. A., & Kuhl, D. E. (1993). Automated detection of the intercommisural line for stereotactic localization of functional brain images. *Journal of Nuclear Medicine*, **34**, 322-329.

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

Nosofsky, R. M. (1992). Exemplars, prototypes, and similarity rules. In A. F. Healy, S. M. Kosslyn, & R. M. Shiffrin (Eds.), *Essays in honor of William K. Estes: Vol. 1. From learning theory to connectionist theory* (pp. 149-167). Hillsdale, NJ: Erlbaum.

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

Nyberg, L., Tulving, E., Habib, R., Nilsson, L., Kapur, S., Houle, S., Cabeza, R., & McIntosh, A. R. (1995). Functional brain maps of retrieval mode and recovery of episodic information. *NeuroReport*, **7**, 249-252.

Pickering, A. D. (1997). New approaches to the study of amnesic patients: What can a neurofunctional philosophy and neural network methods offer? *Memory*, **5**, 255-300.

Poldrack, R. A., Prabhakaran, V., Seger, C. A., & Gabrieli, J. D. E. (1999). Striatal activation during acquisition of a cognitive skill. *Neuropsychology*, **13**, 564-574.

Prabhakaran, V., Smith, J. A. L., Desmond, J. E., Glover, G. H., & Gabrieli, J. D. E. (1997). Neural substrates of fluid reasoning: An fMRI study of neocortical activation during performance on the Raven's Progressive Matrices Test. *Cognitive Psychology*, **33**, 43-63.

Rao, S. M., Bobholz, J. A., Hammeke, T. A., Rosen, A. C., Woodley, S. J., Cunningham, J. M., Cox, R. W., Stein, E. A., & Binder, J. R. (1997). Functional MRI evidence for subcortical participation in conceptual reasoning skills. *NeuroReport*, **8**, 1987-1993.

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

Rugg, M. D., Fletcher, P. C., Frith, C. D., Frambach, M., & Dolan, R. J. (1996). Differential activation of the prefrontal cortex in successful and unsuccessful memory retrieval. *Brain*, **119**, 2073-2083.

Smith, E. E., Patalano, A. L., & Jonides, J. (1998). Alternative strategies of categorization. *Cognition*, **65**, 167-196.

Smith, E. E., & Sloman, S. A. (1994). Similarity- versus rule-based categorization. *Memory & Cognition*, **22**, 377-386.

Smith J. D., & Kemler, D. G. (1984). Overall similarity in adults' classification: The child in all of us. *Journal of Experimental Psychology: General*, **113**, 137-159.

Worsley, J. J., Evans, A. C., Marrett, S., & Neelin, P. (1992). A three-dimensional statistical analysis for CBF activation studies in the human brain. *Journal of Cerebral Blood Flow & Metabolism*, **12**, 900-918.

NOTES

1. The terms "strategy" and "procedure" will be used interchangeably. In no case do we mean to assume deliberate choice of strategy.

2. This paper focuses on the categorization of visual stimuli. Details of the schematic models and corresponding brain regions could, however, be straightforwardly adjusted for other stimulus modalities.

3. Other cognitive processes, such as working memory, may become more heavily involved when highly similar retrieved exemplars suggest conflicting category assignments, but this more complex case will not be investigated here.

4. More specifically, all the rule participants were able to recall the entire rule and reported checking for each rule-related feature either until three were found or until it became clear that three could not be found. The exemplar participants reported memorizing examples and categorizing by assigning items to the same category as the most similar known exemplar.

5. A voxel is a three-dimensional pixel.

(Manuscript received February 23, 2001;
revision accepted for publication July 17, 2001.)