Category Learning in Rhesus Monkeys: A Study of the Shepard, Hovland, and Jenkins (1961) Tasks

3 authors, including:

John Paul Minda
The University of Western Ontario

35 PUBLICATIONS   **1,168** CITATIONS

SEE PROFILE

David Washburn
Georgia State University

145 PUBLICATIONS   **2,905** CITATIONS

SEE PROFILE

Category Learning in Rhesus Monkeys: A Study of the Shepard, Hovland, and Jenkins (1961) Tasks

J. David Smith
University at Buffalo, The State University of New York

John Paul Minda
University of Western Ontario

David A. Washburn
Georgia State University

In influential research, R. N. Shepard, C. I. Hovland, and H. M. Jenkins (1961) surveyed humans' categorization abilities using tasks based in rules, exclusive-or (XOR) relations, and exemplar memorization. Humans' performance was poorly predicted by cue-conditioning or stimulus-generalization theories, causing Shepard et al. to describe it in terms of hypothesis selection and rule application that were possibly supported by verbal mediation. The authors of the current article surveyed monkeys' categorization abilities similarly. Monkeys, like humans, found category tasks with a single relevant dimension the easiest and perceptually chaotic tasks requiring exemplar memorization the most difficult. Monkeys, unlike humans, found tasks based in XOR relations very difficult. The authors discuss the character and basis of the species difference in categorization and consider whether monkeys are the generalization-based cognitive system that humans are not.

Forming psychological equivalence classes is an important cognitive function for humans and nonhuman animals (hereafter referred to as animals). For example, predator categories that allow avoidance or defense responses when appropriate can confer a survival advantage. Vervet monkeys even have call signs for their predator categories (Cheney & Seyfarth, 1990; Struhsaker, 1967). Given the importance of categorization, researchers have focused on animals' categorization abilities, especially those of birds (Bhatt, Wasserman, Reynolds, & Knauss, 1988; Cerella, 1979; Cook, 2001; Herrnstein, Loveland, & Cable, 1976; Huber, 2001; Huber & Lenz, 1993; Lea & Ryan, 1990; Pearce, 1988; Vaughn & Greene, 1984; Wasserman, Kiedinger, & Bhatt, 1988) but also those of primates (D'Amato & Van Sant, 1988; Medin & Dewey, 1984; W. A. Roberts & Mazmanian, 1988; Schrier & Brady, 1987). Our research contributes to this literature an extensive survey of monkeys' categorization and a comparison with humans performing the same tasks.

Four theoretical questions and their associated category tasks underlie our survey of monkeys' categorization. First, researchers have asked whether participants isolate the relevant cue in a categorization task slowly through associative and conditioning processes or suddenly through rule discovery and insight (Harlow, 1949; Krechevsky, 1932; Trabasso & Bower, 1968; Zeaman & House, 1963). This research has relied on category tasks that are solved by focusing on a single diagnostic feature. This kind of task is called a *multidimensional discrimination task* or a *criterial-attribute task* (Smith, Tracy, & Murray, 1993), and it has a long history.

Second, researchers have evaluated participants' ability to learn ill-defined or polymorphous concepts. This research relies on family-resemblance categorization tasks in which no one dimension is sufficient to diagnose category membership but in which several dimensions very likely carry useful category information. Thus category members have a perceptual similarity or a family resemblance to one another. This research is illustrated by many studies of pigeons' categorization (e.g., Huber, 2001; Huber & Lenz, 1993; Jitsumori, 1996; von Fersen & Lea, 1990).

Third, researchers have evaluated participants' capacity to memorize the individual exemplars in category tasks and their correct discriminative responses (Astley & Wasserman, 1992; Pearce, 1987, 1994a, 1994b). This research has relied on perceptually unstructured categories in which the category members have no special similarity to one another. Under these conditions, the participant cannot solve the task using rules or family-resemblance similarity and so must depend on exemplar–response memorization instead. This research is illustrated by studies on pigeons' learning of line-squiggle or other pseudocategories (Vaughn & Greene, 1984; Wasserman et al., 1988). Pigeons have highly effective exemplar memories. Less is known about monkeys' capacity for exemplar memorization, and this helped motivate the choice of species here.

Fourth, researchers have asked whether animals can solve discrimination problems that embody the exclusive-or (XOR) problem (Flagg, 1974; Medin, 1975). For example, suppose that stimuli can be small or large and black or white, and that small white and large black stimuli belong to Category A, whereas large white and small black stimuli belong to Category B. Neither small, large, white, nor black works as a cue to solve the problem—all are uncorrelated with the correct response. (This is why single-layer neural networks cannot solve this problem.) Instead, the problem must be solved through configural learning about the relations between stimulus cues. At present, too little is known about animals' capacity for solving XOR problems (but see Chase & Heinemann, 2001; Pearce, 1994a, 1994b), especially at the higher dimensionalities typical of categorization tasks. This capacity and the different capacities of humans and nonhuman primates for learning XOR tasks are especially important in the present article.

Shepard, Hovland, and Jenkins (1961)

While exploring this same range of capacities in humans (i.e., the capacities to learn rules, family resemblances, cue configurations, and exemplars), Shepard et al. (1961) made an influential methodological contribution. They realized that one could survey these capacities by allocating to categories in different ways the same stimuli composed of highly discriminable dimensional values while also always giving participants the same-sized categories and the same response complexity. Thus they realized that many aspects of the situation could be controlled so that one could ask comparably how well participants solve different types of category problems. This approach helped address a lasting concern with categorization research on humans—that is, the lack of consistency and control over stimuli, category size, and response complexity in different tasks. This approach has the same potential to be useful in research on categorization by animals. The present research contributes to the comparative literature by testing animals for the first time using the tasks and framework of Shepard et al.

To gain this control, Shepard et al. (1961) used three binary dimensions, which we illustrate using the three dimensions—size, brightness, and shape—that ground the present research. These dimensions yield eight stimuli that are coded on the three binary dimensions as shown in Table 1. These stimuli can be arranged into category tasks in many ways, although all of these arrangements are assignable to the six logical task types that are illustrated in Figure 1 and discussed now.

Table 1
Abstract Stimulus Codings and Concrete Stimulus Descriptions for the Stimuli Used in the Present Version of the Shepard, Hovland, and Jenkins (1961) Tasks

Stimulus	Stimulus coding	Stimulus description
1	0 0 0	Small white square
2	0 0 1	Small white triangle
3	0 1 0	Small black square
4	0 1 1	Small black triangle
5	1 0 0	Large white square
6	1 0 1	Large white triangle
7	1 1 0	Large black square
8	1 1 1	Large black triangle

The Type I task has a 100% diagnostic cue: color in the figure. It is a criterial-attribute task. The Type II task instantiates the XOR problem along two of the three dimensions. Neither the color nor the shape information alone is helpful in categorization. But the relation between the color and the shape of the stimulus is helpful. The positively correlated dimensional values 11 and 00 along these dimensions produce the Category A stimuli (black triangles and white squares, respectively). The negatively correlated dimensional values 01 and 10 produce the Category B stimuli (white triangles and black squares, respectively). For this reason, the Type II task is sometimes called a *correlated-features task*. The Type III and Type V tasks are rule-plus-exception tasks, with the rule leaving an exception item that requires additional cognitive processing to master (e.g., exemplar memorization). The Type IV task is a family-resemblance task (although it can also be interpreted as a rule-plus-exception task). The prototypes of each category—a large black triangle for Category A and a small white square for Category B—are joined by the stimuli that share two of three features with their prototype. The Type VI task is chaotic perceptually because its category members have no family resemblance. Each category member shares only one feature with its category mates but two features with several members of the other category. Neither rules nor similarity-based categorization strategies will help performance. Shepard et al. (1961) understood that this category structure would force the participant to rote memorize the category label for each stimulus. Supporting this idea, a current formal model of category learning (SUSTAIN, or supervised and unsupervised stratified adaptive incremental network; Love & Medin, 1998) solves the Type VI problem by devoting one network node to each of the eight exemplars in the task. The Shepard et al. tasks have been highly influential in the categorization literature, receiving additional empirical, formal analytic, and theoretical attention (Feldman, 2000; Kruschke, 1992; Love & Medin, 1998; Nosofsky, Gluck, Palmeri, McKinley, & Glauthier, 1994).

Humans' Performance in the Shepard et al. (1961) Tasks

Shepard et al.'s (1961) empirical and theoretical purpose was to ask whether humans' performance in these six task types was compatible with theories that linked category learning to cue conditioning or stimulus generalization. According to the cue-conditioning hypothesis, a stimulus is a set of cues that can be independently conditioned by trial outcomes. In a Type I category task, for example, Response A might be reinforced only for black stimuli, strengthening an associative connection between the cue black and Response A (whereas white came to control Response B). Noncolor cues, not differentially reinforced in this task, would not exert stimulus control. Under this hypothesis, the difficulty of the categorization task should increase as cue-conditioning processes are less sufficient for complete learning, leaving more supplementary cognitive work to be done.

According to the stimulus-generalization hypothesis, solving a category task is the process of correctly mapping responses to stimuli. Stimulus generalization will play a role in this process because one will confuse similar stimuli and cross-generalize their responses. As that confusion occurs within category or between categories, the mistaken response will still be correct or will be incorrect, respectively. Therefore, category tasks should be harder to master as stimuli become less similar within categories (making
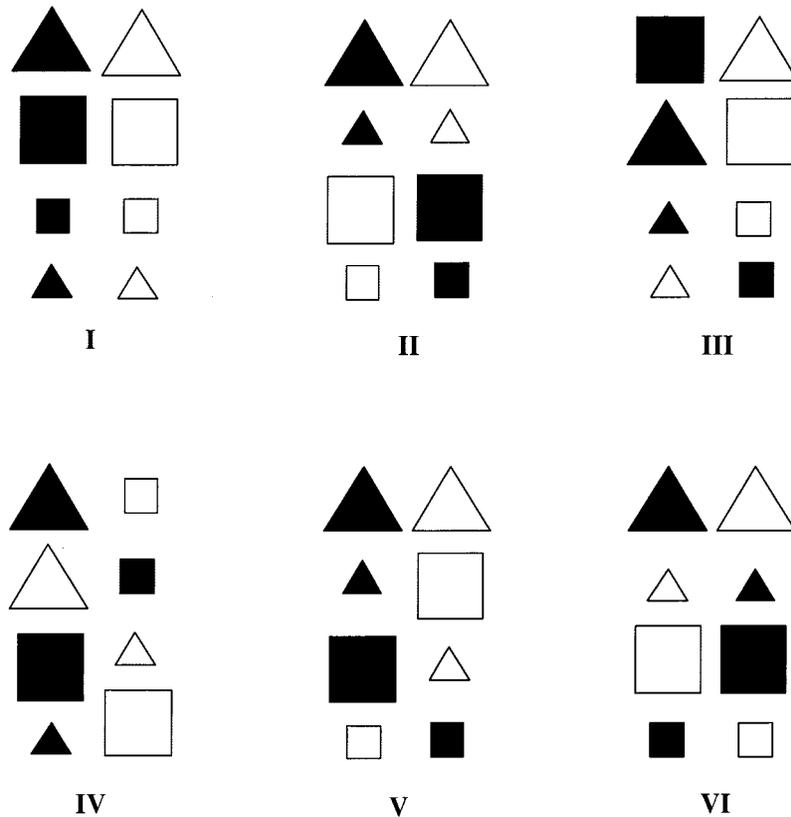
Figure 1.   Examples of the six types of category task used by Shepard et al. (1961) and in the present research.

harmless stimulus-generalization errors less frequent) and more similar between categories (making harmful stimulus-generalization errors more frequent).

Assuming that the stimulus-generalization hypothesis was true, Shepard et al. (1961) predicted the number of errors that participants should make in each task. They based these predictions on the stimulus confusions participants made in an identification task that required eight different labels to be given to eight stimuli. They mapped for each task type which confusions would occur within category (sustaining performance) or between categories (hurting performance) and thus found the predicted errors for each. Figure 2A shows their well-known comparison of these predictions to humans' performance. The points would fall on the straight line if observation fit prediction perfectly. Figure 2B shows the same comparison using a measure of within-category coherence and between-category differentiation that is more easily applied across species (because the elaborate set of identification data is not needed).[1]

Shepard et al. (1961) concluded for several reasons that generalization and conditioning processes were not what underlay humans' category learning in these tasks. Participants learned these category tasks too quickly. For example, in the Type I tasks, they made only about 10 errors in learning instead of the 50 predicted (Figure 2A). For another example, participants learned Type II tasks far faster than the generalization hypothesis allowed and in the wrong rank order of difficulty. Type II was the second easiest task to learn—it should have been the second most difficult. This

shows that humans master XOR tasks surprisingly well, and it raises the question of how nonhuman primates will fare in this kind of task.

On the basis of these findings, Shepard et al. (1961) concluded that humans "are no longer regarded as passively confronting one population of cues after another while a certain crucial subset becomes gradually connected to the correct response" (p. 33). That is, they concluded that humans were not simply obedient to the forces of conditioning and stimulus generalization in these tasks. Instead, they came to regard humans as "abstracting (or attending to) dimensions, and then formulating and testing rules about how the values on those dimensions combine and interact to determine

---

[1] To derive these measures of category coherence and differentiation, we assumed that the similarity between stimuli having no, one, two, or three nonshared features was e^0, e^{-1}, e^{-2}, and e^{-3}, respectively, or 1.0, .37, .135, and .049. This followed the commonly used multiplicative (i.e., exponentially decaying) metric of similarity (Medin & Schaffer, 1978; Nosofsky, 1986). We then measured the summed similarity of each item to the four members of the opposing category and to the four members of its own category (including itself). Dividing these two summed similarities produced a measure of category structure that is commonly called the *structural ratio*. A ratio of 1.0 means that the between-category similarity is as large as within-category similarity and thus that the category task is perceptually chaotic or unstructured. In contrast, a ratio of .37 means that the within-category similarity was much larger and thus that the categories had good perceptual coherence and sharp differentiation from each other.
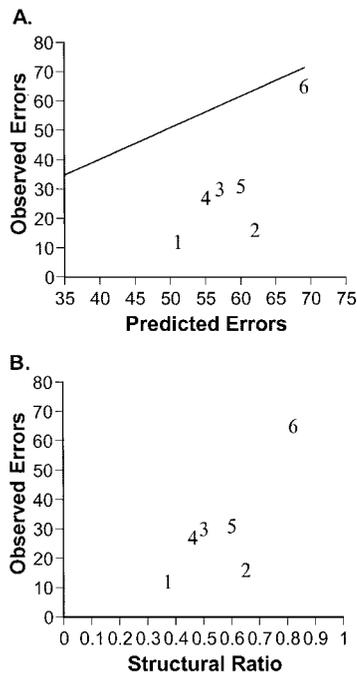
A.

B.

Figure 2. A: The number of errors made by humans as they learned the six types of category task (identified here in arabic rather than roman numerals), plotted against the number of errors predicted by an intuitive similarity-based theory (Shepard et al., 1961). The prediction was based on the number of interstimulus confusions participants made when given an identification task that required them to learn eight separate labels for eight stimuli. B: The number of errors made by humans as they learned the six types of category task, plotted against a measure of the category coherence and differentiation available within each task type. The derivation of this measure is described in Footnote 1.

which classificatory response will be correct" (p. 33). That is, they concluded that humans were active hypothesis testers and rule users. The Shepard et al. study is one of the incisive theoretical expositions in the literature of cognitive science, especially coming as it did in the behaviorist theoretical climate and minimalist computational environment of 1961.

Rationale for the Present Study

This discussion makes clear the reasons for repeating Shepard et al.'s (1961) study with nonhuman primates. These tasks would be an informative survey of categorization by monkeys, as they were for humans. This survey would have the advantages that the size of the categories and the dimensional structure of the stimulus materials would both be controlled across task types. An additional benefit is that we could give the animals the same tasks and stimuli that the humans receive and compare the categorization capacities and strategies of the two species more directly than is usually possible.

This comparison could shed an important light on monkeys' categorization system. Sometimes monkeys have seemed capable of rule use and sudden set switching (Harlow, 1949). Sometimes monkeys have seemed to be more mediational than merely associative in their cognition (Rumbaugh & Pate, 1984). Neither of

these influential paradigms involved category learning, so the extension of these ideas to categorization remains important. Monkeys might transcend conditioning or generalization processes in the Shepard et al. (1961) tasks, too, and show something like the human pattern. Alternatively, the tasks could show that monkeys are associative and generalization based in their categorization performance. If so, they would show poor Type II performance and overall a linear relation between performance and category structure that is unlike the human data pattern.

This possibility bears on an additional comparative contribution that the present research could make. Shepard et al. (1961) sought in humans an associative species that categorized through the processes of stimulus generalization. They found that in their procedures humans are not that species. But an interesting question has remained for 40 years: Which species do instantiate this associative categorization system? If monkeys do, this raises further comparative questions regarding why humans are as they are, why monkeys are limited in this way, and when in evolution these different cognitive organizations emerged.

Method

Monkeys

Participants. Four male rhesus monkeys (*Macaca mulatta*) participated. Hank and Gale were 16 years old at testing; Murph and Lou were 7 years old. The animals had been trained, as described elsewhere, to respond to computer-graphic stimuli by manipulating a joystick (Rumbaugh, Richardson, Washburn, Savage-Rumbaugh, & Hopkins, 1989; Washburn & Rumbaugh, 1991b). The monkeys had also been tested with other computerized tasks but not with a category-learning task as here. The monkeys were tested in their home cages with ad-lib access to the test apparatus, working or resting as they chose. They were not deprived of food or reduced in weight for the purposes of testing, and they had continuous access to water.

Apparatus. Monkeys were tested—as described by Rumbaugh et al. (1989)—with a computerized system consisting of a 486-based computer, an analog joystick, a color monitor, and a pellet dispenser. Monkeys could reach through the mesh of their home cages to manipulate the joystick, which resulted in isomorphic movements of a computer-graphic cursor on the screen. Bringing the cursor into contact with appropriate stimuli earned 97-mg fruit-flavored chow pellets (P. J. Noyes, Lancaster, NH), which were delivered through a Gerbrands 5120 dispenser interfaced to the computer with a relay box and output board (Keithley PIO-12 & ERA-01; Cleveland, OH).

Stimuli. We used the dimensions of variation shown in Figure 1. The shapes were white or dark gray and small or large triangles or squares presented on a colored background. The small and large squares were 40 and 80 screen pixels on a side (about 1.5 and 3.2 cm on the monkeys' 13-inch monitors). The small and large triangles were 51 and 94 screen pixels on a side (about 2.0 and 3.8 cm). Because squares are subjectively larger figures, the side measurements for the triangles were adjusted in pilot studies so that the triangles and squares seemed to be the same size to humans.

Tasks. We chose six versions of each of the six task types to build a battery of 36 tasks. Type I tasks varied in the criterial dimension that allowed the correct categorization of all eight stimuli. Given three possible criterial features and two possible assignments of stimulus groups to Category A or Category B, there were six of these tasks in all. We included all six in our battery, with the features small, large, white, black, square, and triangle criterially associated with Category A.

Type II tasks varied in which dimensions correlated. Given three possible pairs of correlated dimensions (size–color, size–shape, color–shape)

and two possible assignments of stimulus groups to Category A or Category B (i.e., placing positively correlated stimuli in Category A or Category B), there were six of these tasks to include in our battery.

Type III and Type V tasks can be built in 24 ways. Each of the eight stimuli can be the first member of Category A, and then the task can have three different dimensional organizations that dictate the placement of the other stimuli. For consistency's sake, given the use of reversed tasks for Type I and Type II, we chose three Type III and three Type V tasks at random and then included these and their category-reversed versions in our battery.

Type IV tasks can be built in eight ways. Each of the eight stimuli can become the prototype of Category A (with its complementary stimulus being the prototype of Category B and the placement of all eight stimuli dictated thereby). We chose three Type IV tasks at random and then used these and their category-reversed versions in our battery.

Type VI tasks can be built in two ways—there is just one possible partition of the eight stimuli, but one can reverse the category labels associated with the subsets of four stimuli. We used three repetitions of the two possible versions in our battery.

Table 2 shows the Category A and Category B stimuli for the 36 tasks. The stimulus numbers correspond to the stimulus descriptions in Table 1.

Table 2
Task Type and Stimulus Specifications by Category for the 36 Category Problems Included in the Present Version of the Shepard, Hovland, and Jenkins (1961) Tasks

Task	Type	Category A stimuli	Category B stimuli
1	I	1 2 3 4	5 6 7 8
2	I	1 2 5 6	3 4 7 8
3	I	1 3 5 7	2 4 6 8
4	I	5 6 7 8	1 2 3 4
5	I	3 4 7 8	1 2 5 6
6	I	2 4 6 8	1 3 5 7
7	II	1 2 7 8	3 4 5 6
8	II	1 3 6 8	2 4 5 7
9	II	1 5 4 8	2 6 3 7
10	II	3 4 5 6	1 2 7 8
11	II	2 4 5 7	1 3 6 8
12	II	2 6 3 7	1 5 4 8
13	III	1 2 3 6	8 7 4 5
14	III	5 6 7 2	4 3 8 1
15	III	4 3 2 7	5 6 1 8
16	III	8 7 4 5	1 2 3 6
17	III	4 3 8 1	5 6 7 2
18	III	5 6 1 8	4 3 2 7
19	IV	2 6 4 1	7 3 5 8
20	IV	3 7 1 4	6 2 8 5
21	IV	4 8 2 3	5 1 7 6
22	IV	7 3 5 8	2 6 4 1
23	IV	6 2 8 5	3 7 1 4
24	IV	5 1 7 6	4 8 2 3
25	V	1 2 3 8	5 6 7 4
26	V	2 1 4 7	6 5 8 3
27	V	7 8 5 2	3 4 1 6
28	V	5 6 7 4	1 2 3 8
29	V	6 5 8 3	2 1 4 7
30	V	3 4 1 6	7 8 5 2
31	VI	1 4 6 7	2 3 5 8
32	VI	2 3 5 8	1 4 6 7
33	VI	1 4 6 7	2 3 5 8
34	VI	2 3 5 8	1 4 6 7
35	VI	1 4 6 7	2 3 5 8
36	VI	2 3 5 8	1 4 6 7

Order of tasks. These 36 tasks were arranged into a random order such that each six tasks represented one permutation of the six task types. Two monkeys received this random order and two the opposite order.

Procedure. One task was given per day. Each day the background screen color changed to remind the animal that the rules and stimulus assignments were now different. These color changes were made in a random permutation of seven colors so that screen colors would not coordinate with the task-type changes that were made in a random permutation of six. The animals generally received 2,400 trials of each task. Technical problems disrupted 2 sessions of the 144 (4 monkeys × 36 tasks) when the monkeys were between 2,000 and 2,400 trials, so we chose to analyze the first 2,000 trials of each session.

In all, the four animals completed nearly 350,000 trials, of which 288,000 trials are analyzed here (72,000 per animal). Each trial operated as follows. On the screen were a large X and O to the left and right, respectively. The to-be-categorized object was centered between these response destinations. The animal used the joystick to move the to-be-categorized object to the X or the O. Trials were presented in random permutations of the eight shapes, but nothing in the appearance of the task indicated this blocked structure. The animals received sound feedback and a food pellet for each correct response and a time-out period of 10 s signaled by a buzzing sound for each incorrect response. The buzzing sound began immediately with the completion of the incorrect response and occurred with the stimulus object and the wrong response choice visible. In both cases, the next trial followed immediately. The animals were able to complete 10 trials in about 40 s, performing as they did at about 80% correct overall. This could have produced about 15 trials per minute and a session length for 2,400 trials of around 3 hr. However, the animals were free to drink, rest, or take any other off-task time they chose. Thus the actual session lengths were longer (4–6 hr).

Humans

Participants. Forty-seven undergraduates participated. They received course credit and competed for cash prizes, which were given to the top point scorers.

Procedure. The procedure already described was basically followed. Instead of receiving food rewards, humans saw their points continuously updated on the screen. Humans received a 1-s time-out and no points for incorrect responses. They responded by pressing keys labeled *1* and *2* that corresponded with the screen positions of a 1 and a 2 that were placed just like the X and O characters used for the monkeys. The stimulus dimensions and specifications were those already described.

Each human received one category task of each type. The types were presented in one of six permutations. Over each 6 participants, each type appeared equally often in each serial position within the session. The specific version of the type of task—chosen from among the six of each type that were given to the animals—was also assigned using one of six permutations. The tasks were given to the participants consecutively, with different screen colors again reminding the participants that the stimulus assignments had changed. Humans were given 192 trials (24 blocks of the eight stimuli) for each task.

Human participants received these instructions:

> This experiment is about grouping and classification. You will see shapes that can be classified as "Category 1" shapes or as "Category 2" shapes. Your job is to look carefully at each shape and decide if it belongs to Category 1 or Category 2. Type the key labeled "1" on the keyboard if you think it is a Category 1 shape. Type the key labeled "2" if you think it is a Category 2 shape. If you choose correctly you will hear a Whoop sound; if you choose incorrectly you will hear a low, buzzing sound. At first the task will seem quite difficult, but with time and practice you should be able to answer correctly. This task has no tricks or surprises. There are always correct answers for what categories the shapes belong in, and once you learn the right catego-

ries or rules they will work to classify the shapes for the whole task you are on.

After you have done a series of trials on one task, you will see the background screen color change. This means that you are being given a new task. At this point you must learn again from the start which shapes are in Category 1 or Category 2, because the rules and relevant dimensions will be decided independently for the new task. You will complete six tasks in all, and you will have a maximum of 192 trials on each. Try to get as many trials as you can correct on each of the six tasks. The person in this week's experiment who gets the highest percentage correct over all six tasks will receive a $20 prize. The second place winner will receive a $10 prize. The third place winner will receive a $5 prize. So the experiment is fair to everybody, PLEASE USE NO NOTES OR WRITING OF ANY KIND. Just learn the tasks in your own mind ONLY.

There are differences to consider between the human and animal procedures. As is traditionally done and as was done by Shepard et al. (1961), humans were given instructions about learning categories. Humans were also told that a new screen color meant a new task. Humans received all their tasks in one day instead of on successive days—this might have affected the memory processes used or increased memory interference across tasks. Humans also received fewer trials in each task because they learned so quickly. These procedural differences involve almost universal aspects of procedure in the human literature (e.g., instructions, fewer trials, single-day sessions). We consider their implications in the discussion and suggest lines of research that follow from them. Here we simply note that our procedures created the closest existing comparison between human and monkey categorization but not a perfectly exact one.

Results

Humans

We summarize first the results from humans to confirm that our stimuli and methods reproduced Shepard et al.'s (1961) results.

Data summary. Table 3 shows the percentage of correct responses for humans during each odd eight-trial block and across all 24 trial blocks. Performance is summarized across participants and task types (Row 1) and across participants within task type (Rows 2–7). We will point to some notable patterns in this summary table.

Overall performance levels. Row 1 of Table 3 shows that the humans achieved 78.2% correct over all 54,144 trials (24 trial blocks × 8 trials per block × 6 tasks × 47 participants), a good level of performance given that this includes all stages of acquisition and given that all tasks ended after 192 trials (one tenth of the trials the monkeys received). There were considerable individ-

ual differences in performance, though, with individual humans averaging as high as 94% and as low as 63% correct overall.

Learning within tasks. Row 1 also shows the overall course of acquisition. Humans achieved most of their learning early on and then showed slight improvement later on. On average, in their first, seventh, and last blocks, respectively, humans were 50.9%, 75.3%, and 86.2% correct.

Order of task difficulty. Shepard et al. (1961) found that humans showed this rank order of task difficulty: I < II < III = IV = V < VI (see Figure 2). Nosofsky et al. (1994) replicated this rank order of difficulty. Love (2002) did as well for a subset of the tasks (I, II, IV, and VI). Rows 2–7 of Table 3 show that we obtained the same rank order. Figure 3 summarizes these four results to show archivally the consistency of this rank order. Using the logarithmic transformation of errors to criterion that Shepard et al. used, we confirmed in our data that the Type I task was significantly easier than the Type II task, $Ms = 2.33$ ($SD = 1.178$) and 4.52 ($SD = 1.564$), respectively, $t(92) = -6.96$, $p < .01$; Type II was easier than Types III, IV, and V, $Ms = 4.52$ ($SD = 1.564$) and 5.11 ($SD = 1.280$), respectively, $t(186) = -2.59$, $p < .01$; and Types III, IV, and V were easier than Type VI, $Ms = 5.11$ ($SD = 1.280$) and 5.61 ($SD = 1.211$), respectively, $t(186) = -2.35$, $p < .01$. (Means of 2.0, 3.0, 4.0, and 5.0 would correspond to 4, 8, 16, and 32 errors to criterion, respectively.) The cross-task patterning may have been slightly less strong in our case than in Shepard et al.'s, perhaps because they used more highly selected participants than we did or because they used different stimulus dimensions. Our patterning may have been stronger in one respect than that found by Love (2002)—he was unable to statistically distinguish Type II and Type IV performance. Love also discussed the factors that can affect the strength of the order effects across the Shepard et al. tasks.

Rule discovery in Type I tasks. The Type I or criterial-attribute task is useful because it directly evaluates how quickly and flexibly organisms isolate and attend to the single relevant dimension. If participants realize the correct dimensional rule, they should strike the Type I task's solution suddenly and show a near-vertical segment in their acquisition curve. But if the associative strength between stimulus cues and categorization responses increases gradually through conditioning processes, participants should approach the solution more slowly with gentler sloped acquisition curves. These possibilities were critical to Zeaman and House's (1963) analysis of the discrimination learning of children with and

Table 3

Percentage of Correct Responses by Humans During Each Odd Eight-Trial Block (Columns 1–23) and Across All 24 Trial Blocks (Column All)

Row	Type	1	3	5	7	9	11	13	15	17	19	21	23	All
1		50.9	67.6	71.1	75.3	78.1	81.5	81.8	82.6	83.9	84.9	85.9	85.1	78.2
2	I	64.4	94.1	92.6	96.5	95.7	98.1	95.5	96.5	98.7	98.9	97.6	96.8	94.6
3	II	49.7	68.1	74.7	77.4	83.2	82.4	84.8	85.4	86.7	86.7	89.9	85.6	80.3
4	III	52.9	63.8	69.4	74.2	76.9	80.9	80.9	80.1	85.6	83.0	84.0	84.0	77.1
5	IV	47.9	62.5	66.2	67.6	71.8	77.7	77.1	81.1	76.3	83.8	83.8	85.4	74.7
6	V	48.1	60.4	64.1	68.6	73.9	80.3	83.0	81.6	84.0	84.0	84.6	85.1	75.3
7	VI	42.6	56.9	59.8	67.3	67.0	69.4	69.7	70.7	71.8	72.9	75.3	73.4	67.1

Note. Performance is summarized across six task types (Row 1) and for each type separately (Rows 2–7).
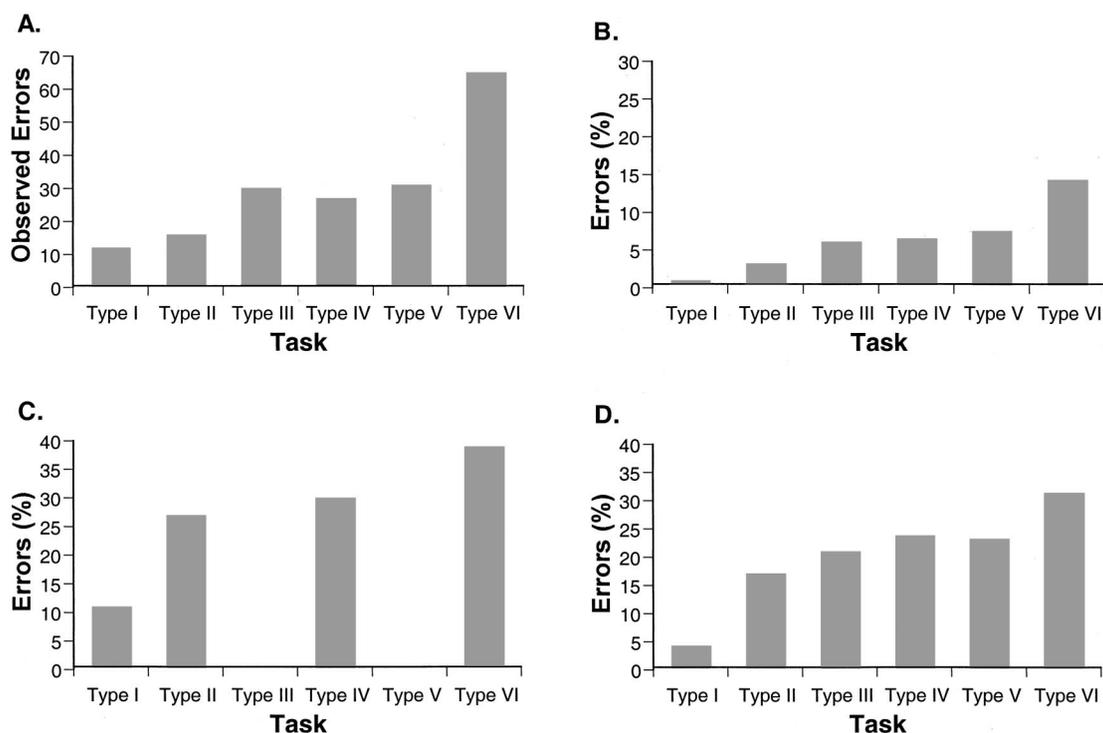
Figure 3. Performance by humans in the Shepard et al. (1961) category tasks in three existing studies and the present study. A: Results from the Shepard et al. study. B: Results from the Nosofsky et al. (1994) study. C: Results from the Love (2002) study. D: Results from the present study.

without mental retardation and to Smith et al.'s (1993) analysis of the category learning of individuals with and without depression.

Figure 4A shows humans' acquisition of Type I tasks over 24 eight-trial blocks. The percentage of correct responses rises rapidly from 64% to 94% in the third block (i.e., after only 16 trials). Humans realize very quickly the appropriate attentional strategy in this task. In fact, Figure 4A underestimates the suddenness of rule discovery. The shape of this group-forward acquisition curve (with the group's performance measured forward from the onset of training) could be the result of discoveries at different blocks for different participants. Rule discovery may be seen more clearly through a backward learning curve that aligns the criterion run of all participants and then summarizes performance one trial block before criterion, two blocks before criterion, and so forth (Hayes, 1953; Smith et al., 1993; Zeaman & House, 1963). In the present case, we found the place where the participant began a run of three perfect blocks and then numbered blocks backward and forward from this. Figure 4B shows the backward learning curve for the humans for the Type I task. (The data from 45 of the 47 participants were included—1 participant never reached criterion and 1 participant never erred and so had no precriterion data.) A *t* test confirmed the significance of the performance change in the transition to criterion, Ms = 75.6 (SD = 15.3) for the block before the criterion run and 100.0 (SD = 0.0) for the criterion run, $t(44) = -10.72$, $p < .01$. Clearly the cognitive reorganization that led to the criterion run was sudden.[2] It also was lasting. An important feature of this backward learning curve is that performance following the criterion run continued at a very high level (96.9%

correct, SD = 6.1, which is still statistically remote from the precriterion performance level, $t[44] = -8.15$, $p < .01$).

We conclude from our analyses of Type I tasks that humans go through a process of rule discovery that leads to their sudden and subsequent correct performance. Thus our results confirm the principal conclusion of Shepard et al. (1961), which was that humans were discovering rules rather than gradually associating responses with cues and rather than assigning category labels in a way that was dependent on stimulus generalization among similar stimuli.

Rule discovery in Type II tasks. We also evaluated the character of humans' rule discovery in Type II (correlated-feature) tasks. The forward learning curve in Figure 4C is gently sloped, which could indicate a gradual learning process, perhaps based in the conditioning of configural cues. Or this curve could be the

---

[2] Notice that the cognitive reorganization that leads to criterion begins from a level of about 75% correct. The strictest interpretation of human rule use would suppose that humans should perform at chance while they are evaluating incorrect hypotheses about the task's solution, and thus the performance jump to criterion would be from a baseline of 50% (e.g., Trabasso & Bower, 1968; Zeaman & House, 1963). However, 40 years of subsequent research and theory make it reasonable to assume that other kinds of learning (e.g., learning the correct responses for a couple of individual exemplars in the task) could raise precriterion learning above the 50% level. In fact, Figures 5 and 8 in Smith et al. (1993, pp. 338 and 343, respectively) show that some other learning process does raise precriterion performance above 50% but that still, as here, there is a sudden increase in performance accuracy when the rule is discovered.
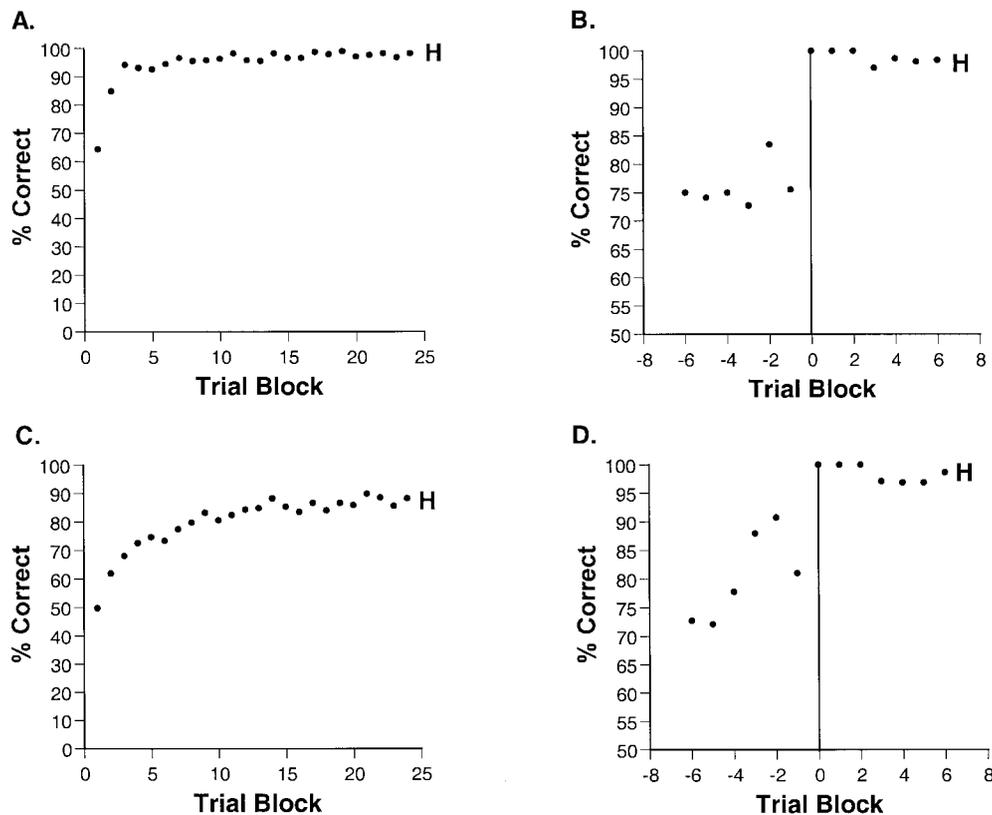
Figure 4. A: Humans' (H) forward learning curve in Type I tasks. Their percentage correct is shown at each eight-trial block. B: Humans' backward learning curve in Type I tasks. Their percentage correct is shown for eight-trial blocks forward and backward from the start of their first run of three consecutive perfect blocks. C: Humans' forward learning curve in Type II tasks. D: Humans' backward learning curve in Type II tasks.

result of sudden insights made by different participants at different trial blocks. The backward learning curve in Figure 4D shows that many humans did reach criterion in the Type II task suddenly, although the transition was more gradual than we saw in Type I task performance. A *t* test confirmed the significance of the performance change in the transition to criterion, $Ms = 81.0$ ($SD = 11.1$) and 100.0 ($SD = 0.0$), $t(30) = -9.49$, $p < .01$. (The data from 31 of 47 participants who reached criterion were included in the backward curve.) For this task, too, the cognitive reorganization that led to the criterion run was rapid. It was also lasting. Performance following the criterion run continued at a high level (i.e., 97.1%, $SD = 5.4$, correct, still statistically remote from the level of precriterion performance, $t[29] = -6.75$, $p < .01$; an additional participant had no postcriterion data). Thus this backward curve supports Shepard et al.'s (1961) theoretical conclusions. Given the replication of Shepard et al.'s results for Type I and Type II tasks in the overall rank order of difficulty and the support for their conclusions provided by backward learning curves, we turned to consider the monkeys' data.

Monkeys

Data summary. Table 4 summarizes the 4 monkeys' performance on six types of category task experienced in six rotations. The data are summarized over all 144 category tasks (Row 1), by

rotation (Rows 2–7), by monkey (Rows 8–11), and by type of category task (Rows 12–17). Each row gives the percentage of correct responses the monkeys achieved during each odd 80-trial block and across all 25 trial blocks. We point to some notable patterns in this summary table.

Row 1 of Table 4 shows that the monkeys achieved 80.2% correct over all 288,000 trials analyzed, a strong level of performance given that this includes all stages of acquisition. Row 1 also shows the course of acquisition. Animals achieved most of their learning early on and less thereafter. In their first, seventh, and last block, respectively, animals were 58.5%, 79.7%, and 83.9% correct on average.

Rows 2–7 show that monkeys learned to learn in these category tasks. These rows summarize performance for the 24 tasks (4 monkeys × 6 task types) sampled in Sessions 1–6 (Row 2), the 24 tasks sampled in Sessions 7–12 (Row 3), and so forth. The animals learned faster across the successive rotations. For Rotations 1 to 6, respectively, the overall percentages of correct responses were 72.7%, 76.0%, 81.9%, 84.2%, 81.1%, and 85.4%.

Rows 8–11 show that there were substantial individual differences in overall performance across the animals. The four monkeys were, respectively, 88.2%, 90.7%, 77.1%, and 64.9% correct overall.

Monkeys' and humans' relative speeds of learning. These data provide one of the closest existing comparisons between human

Table 4
Percentage of Correct Responses by Monkeys During Each Odd 80-Trial Block (Columns 1–25) and Across All 25 Trial Blocks (Column All)

Row	M	T	R	1	3	5	7	9	11	13	15	17	19	21	23	25	All
1				58.5	74.2	78.1	79.7	80.9	81.7	82.7	83.9	83.6	83.2	82.8	83.7	83.9	80.2
2			1	51.9	64.5	69.4	71.0	73.6	74.3	75.4	76.0	75.6	74.5	76.0	79.6	79.6	72.7
3			2	52.6	67.1	72.7	76.0	77.5	79.0	80.1	81.4	80.1	80.2	79.4	77.7	81.6	76.0
4			3	59.4	77.2	80.4	81.9	81.9	83.3	84.8	85.9	86.5	84.2	83.1	84.9	84.4	81.9
5			4	62.0	80.5	82.9	83.7	86.4	85.5	86.1	86.8	86.6	87.8	86.6	87.4	86.7	84.2
6			5	58.8	75.7	81.0	81.9	80.1	83.1	82.9	84.7	83.1	83.1	83.1	83.4	82.0	81.1
7			6	66.7	80.5	82.0	83.9	86.0	85.2	86.8	88.4	87.9	89.2	88.6	89.3	89.1	85.4
8	1			62.4	83.1	87.4	89.7	90.7	90.1	90.8	91.0	91.4	90.9	89.6	89.8	89.7	88.2
9	2			65.6	87.9	90.7	92.6	92.6	92.7	92.3	93.4	93.2	93.0	92.3	91.8	92.8	90.7
10	3			55.5	68.7	73.6	76.0	77.1	78.9	79.4	81.9	81.1	80.7	81.3	82.3	81.6	77.1
11	4			50.7	57.3	60.5	60.6	63.3	65.2	68.2	69.2	68.6	68.0	68.0	71.0	71.3	64.9
12		I		65.1	79.9	84.1	85.5	86.8	87.9	88.6	89.3	88.6	88.8	88.4	91.6	92.7	86.2
13		II		55.6	69.2	74.6	77.3	77.2	79.8	79.0	82.1	81.0	81.8	81.3	82.2	82.1	77.5
14		III		58.7	76.9	81.5	82.7	84.3	82.1	84.9	85.5	84.9	86.0	86.1	85.3	83.4	82.4
15		IV		59.0	76.6	79.6	81.9	83.5	84.2	83.3	84.7	85.6	84.3	84.0	84.2	84.4	81.8
16		V		60.8	74.8	77.4	78.6	80.8	81.8	84.6	85.1	83.3	84.2	84.4	84.0	83.4	80.7
17		VI		52.1	68.1	71.4	72.4	72.9	74.6	75.6	76.7	78.0	73.7	72.8	75.0	77.3	72.8

Note. Performance is summarized in Row 1 over all 144 category problems (4 monkeys × 6 task types × 6 rotations), by rotation (R; Rows 2–7), by monkey (M; Rows 8–11), and by task type (T; Rows 12–17).

and animal categorization. In relating Tables 3 and 4, remember that the trial segments for monkeys (80 trials) contained 10 times as many trials as did the trial blocks (8 trials) for humans. Thus, at corresponding points in the tables, the monkeys had received 10 times the training humans had received. Figure 5 shows the profound species difference in the speed of learning. This difference accords well with existing research. For example, Medin and Dewey (1984) found that monkeys needed about 1,100 trials to reach 90% correct in an eight-stimulus categorization task. In a similar nine-stimulus task, Medin, Altom, and Murphy (1984) found that 81% of humans achieved an errorless 9-trial block within 144 trials.

Order of task difficulty. The main theoretical question motivating this research concerned the rank order of task difficulty the animals would show. According to generalization theories, we
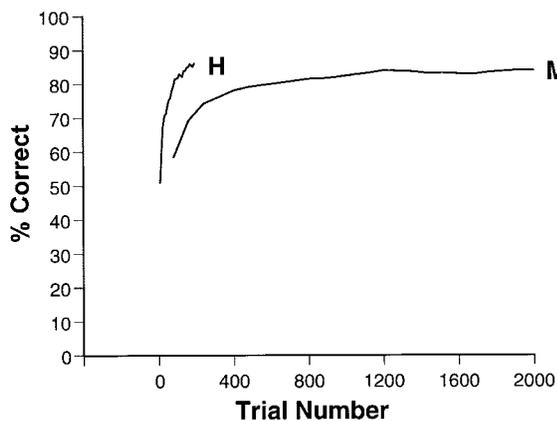

100
90
80 H
70
% Correct 60
50
40
30
20
10
0
0   400   800   1200   1600   2000
Trial Number

Figure 5. A comparison of humans' (H) and monkeys' (M) speed of learning. Human performance is averaged across 6 category tasks for 24 eight-trial blocks. Monkey performance is averaged across 36 category tasks for 25 eighty-trial segments.

would predict—as suggested by Shepard et al. (1961)—a rank ordering of I < III = IV = V < II < VI. However, Shepard et al. found that humans showed the qualitatively different rank ordering of I < II < III = IV = V < VI, suggesting that humans were engaged more in hypothesis testing and rule discovery than in the gradual processes of association and similarity generalization.

Rows 12–17 of Table 4 answer this question for the monkeys. They showed the rank order predicted by similarity-generalization theories. They found Type I tasks to be the easiest, Types III–V next easiest, Type II tasks more difficult, and Type VI tasks most difficult. Testing for the significance of this pattern, we found Type III–V performance to be significantly worse than Type I performance, $Ms = 81.6\%$ ($SD = 11.8$) and 86.2% ($SD = 12.9$), respectively, $t(23) = 2.77$, $p < .05$; Type II performance to be significantly worse than Type III–V performance, $Ms = 77.5\%$ ($SD = 15.2$) and 81.6% ($SD = 11.8$), respectively, $t(23) = 2.63$, $p < .05$; and Type VI performance to be significantly worse than Type II performance, $Ms = 72.8\%$ ($SD = 15.0$) and 77.5% ($SD = 15.2$), respectively, $t(23) = 2.41$, $p < .05$. These paired t tests allowed us to compare performance levels across task types but within monkey and within rotation. This controlled for different animals' overall performance level and for the general improvement that occurred across rotations. The crucial result is that the monkeys found Type II tasks to be the second most difficult, whereas humans found them second easiest. Figure 6 confirms that all four animals showed the identical rank order of difficulty. The pattern held for two strongly performing animals, for a moderately performing animal, and for one animal (Monkey 4) who operationalized "monkeying around."

Researchers of animal behavior may wonder how three kinds of cross-task carryover effects (caused by giving animals successive tasks on successive days) affected the pattern of results across tasks. First, animals might have experienced positive carryover effects. Their learning to learn might have caused the cross-task performance pattern to change through the six rotations of the
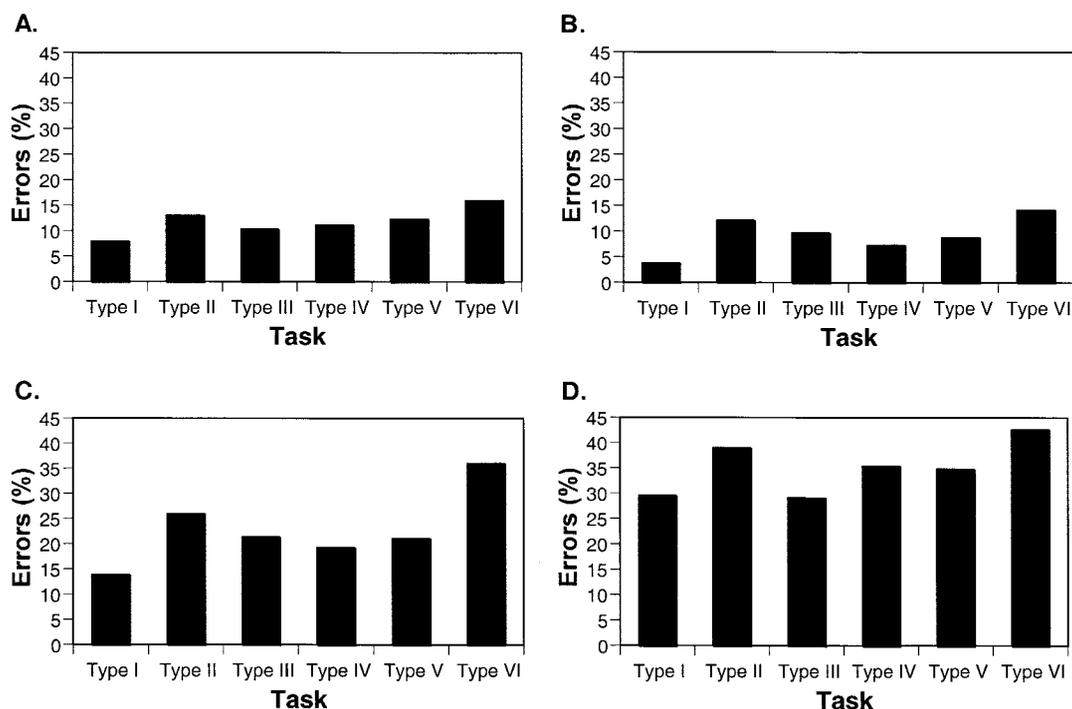
Figure 6. Performance by four individual monkeys in the Shepard et al. (1961) category tasks. A: Performance of Monkey 1. B: Performance of Monkey 2. C: Performance of Monkey 3. D: Performance of Monkey 4.

experiment. This pattern might even have drifted toward the human pattern. To evaluate this drift, we carried out a two-way analysis of variance on the overall percentage of correct responses, with task type and rotation as within-monkey factors. There was no Task × Rotation interaction—the rank order of task difficulty held across rotations.

Second, animals might have experienced negative carryover effects. The same stimuli were used from session to session, but each day some had incongruous category assignments and some had congruous category assignments compared with the assignment from the day before. To explore these effects, we analyzed monkeys' successive task performances by tracking which stimuli carried forward into the next session with the same or a different category assignment. This let us estimate the cross-task pattern of performance twice: separately for the positive-transfer and negative-transfer stimuli. Both estimates produced the same rank order of performance across tasks that we have already described, although the negative-transfer stimuli were performed at somewhat lower levels. Thus negative carryover had a global retarding effect on learning that did not interact with task type to affect the crucial result of the experiment. The global negative-carryover effect, brought about by associative inertia from the previous day, is consistent with the interpretations of monkey performance given below.

Third, animals might have exhibited attentional biases—that is, consistent tendencies to attend to some dimensions and not others. We point out that attentional biases would not have produced the observed rank order of task difficulty. A fully biased animal would be 75% correct on all Type IV tasks (because all cues in that task are 75% diagnostic) but would only be 66% correct on Type I tasks

(100% correct one third of the time, but 50% correct two thirds of the time). This is not the rank order the animals showed. Nonetheless, we examined the data for attentional biases by asking how the animals did on Type I tasks when the different dimensions were relevant and when attention to different dimensions was required. Monkeys 1 and 2 showed excellent performance in Type I tasks no matter which dimension was relevant: Their attention was not biased. Monkeys 3 and 4 did show some bias (Monkey 3 had difficulty attending to shape; Monkey 4 only attended well to size). However, all four animals produced the same rank order of task difficulty, confirming that these biases were not behind the data pattern we observed.

Figure 7A shows the monkeys' errors in each task type plotted against each task's structural ratio (i.e., the measure of perceptual coherence within categories and perceptual differentiation between categories that was already described). Error rates for the monkeys were a straight-line function of the objective difficulty of the task as indexed by perceptual coherence and differentiation. The correlation shown in Figure 7A is .98. Figures 7B–7E show the scatter plots for the individual monkeys. These correlations were .98, .94, .97, and .88. These linear relationships suggest a simple but important theoretical conclusion: Monkeys verge on being the cue-conditioning or similarity-generalization creatures that Shepard et al. (1961) found humans not to be.

Learning in Type I tasks. We created forward and backward learning curves for the monkeys in order to differentiate category learning based in gradually strengthening associations from category learning based in hypothesis testing and rule discovery. Figure 8A shows the forward curve. To make it, we broke the first 1,992 trials of each Type I task into eighty-three 24-trial segments
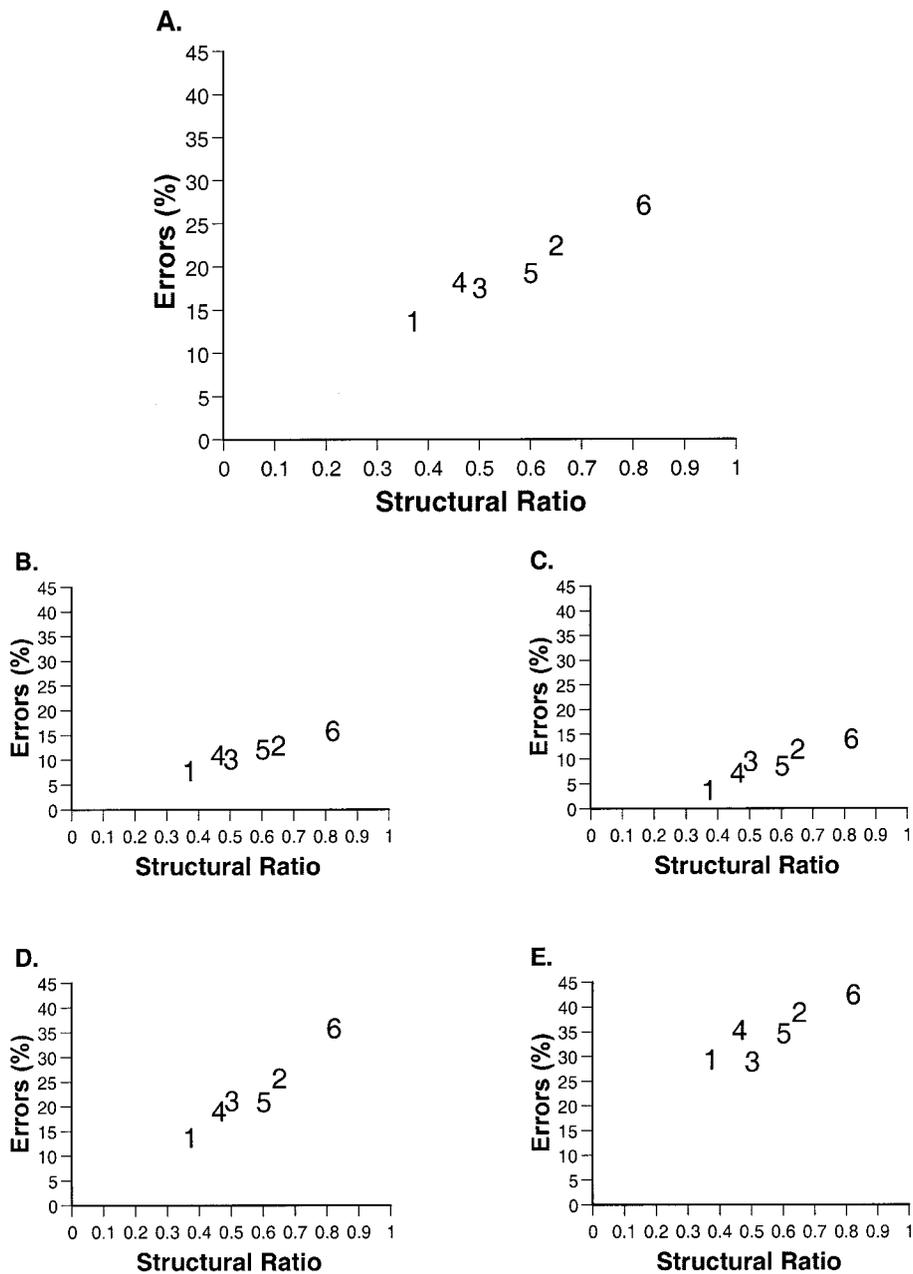
Figure 7.   A: Percentage of errors made by four monkeys as they learned the six types of category task, plotted against a measure (described in Footnote 1) of the category coherence and differentiation available within each task type. B: The scatter plot of Monkey 1's performance. C: The scatter plot of Monkey 2's performance. D: The scatter plot of Monkey 3's performance. E: The scatter plot of Monkey 4's performance.

(each containing three runs through the eight stimuli). An early period of strong improvement extended over about eight 24-trial blocks or 192 trials, during which the monkeys improved from about 55% to 80% correct. (Humans improved to 94% correct in 16 trials.)

Figure 8B shows the backward learning curve anchored to the point of the first perfect 24-trial segment. The data from 18 of 24 Type I tasks were included—Monkey 3 in 2 tasks and Monkey 4 in 4 tasks never achieved a criterion run. The backward curve

shows no point of discovery as it did for humans. In fact, the arrival at criterion was only a chance departure from a gradual acquisition process that continued before and after the criterion block. To see this, note that the criterion level of performance was not sustained. Instead, performance fell back to a level only 2% higher than before criterion (93.1% vs. 91.0%), $t(34) = -0.82$, *ns*. In a case like this, the criterion run is less meaningful and the backward learning curve is less informative, although it was necessary here for comparison to the human case. Neither the forward
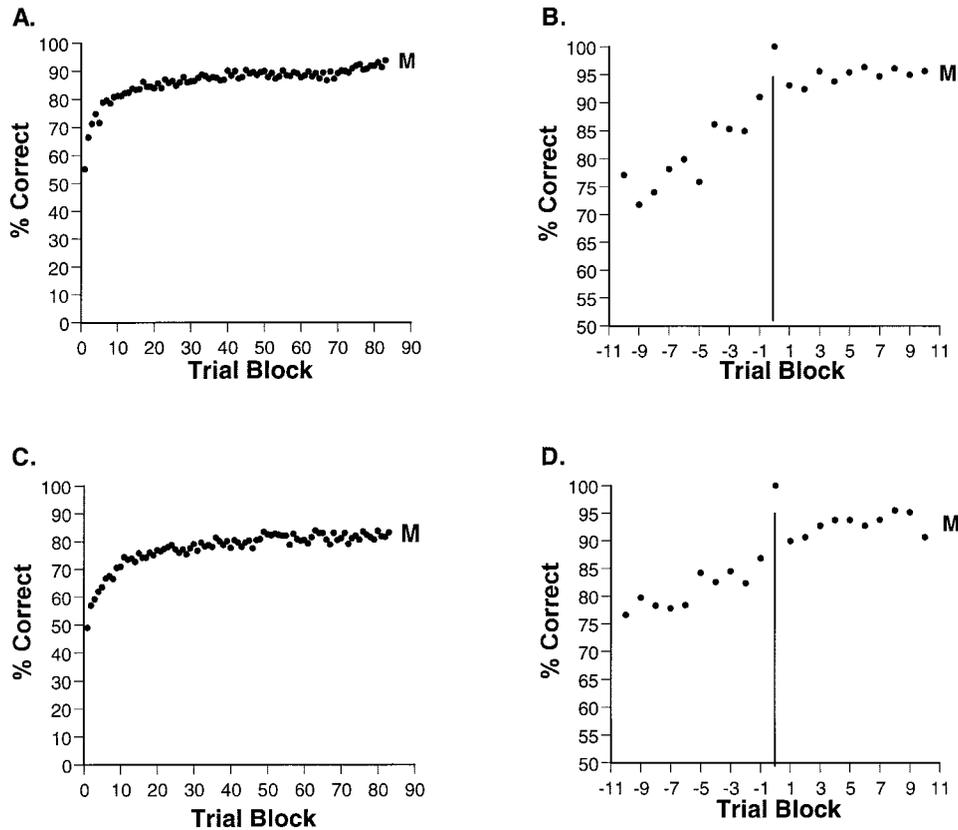
Figure 8. A: Monkeys' (M) forward learning curve in Type I tasks. Their percentage of correct responses is shown at each of eighty-three 24-trial blocks. B: Monkeys' backward learning curve in Type I tasks. Their percentage of correct responses is shown for 24-trial blocks forward and backward from their first perfect 24-trial block. C: Monkeys' forward learning curve in Type II tasks. D: Monkeys' backward learning curve in Type II tasks.

nor the backward curve provides any warrant for the claim that monkeys solve Type I tasks through sudden rule discovery. Rather, both curves imply a gradual process of learning.

Learning in Type II tasks. We also drew forward and backward learning curves for the monkeys' Type II data. The forward curve (Figure 8C) increases more slowly than did the Type I curve and the asymptote is lower than that of the Type I curve. The backward curve (Figure 8D) confirms that a continuous learning process underlay learning. (The data from 12 of 24 Type II tasks were included in the backward curve—Monkeys 1 and 2 in 1 task and Monkeys 3 and 4 in 5 tasks never achieved a perfect 24-trial segment.) Again it is clear that the arrival at criterion was only a chance departure from a gradual learning process that went on before and after that point. Here, too, the criterion level of performance was not sustained. Just after the criterion block, performance fell back to a level that was only 3% higher than before criterion (89.9% vs. 86.8%), $t(22) = -0.88$, *ns*. Thus, rather than a rapid and lasting cognitive reorganization over about 24 trials, as the humans showed, monkeys showed a glacial improvement in performance over hundreds of trials. The Type II tasks underscore the profound species differences in category learning that the Type I tasks revealed. Humans seem to learn some category tasks through a sudden process of rule discovery. Monkeys apparently learn them by gradually associating responses with stimuli through cue conditioning or through slowly sharpening stimulus-generalization processes.

Discussion

Our results from humans confirmed those of Shepard et al. (1961). The observed rank order of difficulty across the tasks was not predicted by assuming cue-conditioning or stimulus-generalization processes. Humans learned Type I tasks more suddenly and Type II tasks more easily than these processes would allow. Backward learning curves isolated the point of rule discovery, extended Shepard et al.'s results, and underscored the theoretical interpretation that the human categorization system transcends associative or similarity-based processes.

In contrast, monkeys' rank order of task difficulty was predicted by assuming cue-conditioning or stimulus-generalization processes. Monkeys learned Type I tasks slowly and Type II tasks with great difficulty. Overall, their performance tracked perfectly the tasks' perceptual coherence and similarity structure (Figure 7), as predicted by an associative or similarity-based view. These results frame our discussion and present a comparative contrast that could ground further research and theoretical development.

A central question concerns the basis of the species differences. Here we consider several possibilities that arise from our theoretical interests. First, Shepard et al. (1961) suggested that humans' language resource and their verbalization of hypotheses selected and rules applied might have produced their data pattern, especially their strong performance on Type I and Type II tasks. Shepard et al. grounded these suggestions by showing that the rank order of task difficulty was—for humans—closely tied to the complexity of the simplest verbal rule for describing the task's solution.[3] Feldman's (2000) related explanation was that humans' rank order of task difficulty in the Shepard et al. tasks was proportional to the tasks' Boolean complexity—that is, to the length of the shortest logical expression that captures each task's solution. Emphasizing further the role of verbal rules and language in humans' performance of these tasks, Shepard et al. also found that as participants learned successive category problems of the same type, the complexity of their stated verbal rule decreased as its parsimony increased. This increase in rule simplicity facilitated learning.

One can see why language would confer a processing advantage by letting humans represent their rules and hypotheses succinctly and store them stably for testing and acceptance or rejection. In fact, the Type I and Type II tasks illustrate the kind of cognitive tasks that would benefit from language coding and that might have exerted a positive pressure toward the language adaptation or toward its cognitive preadaptation. However, as we discuss presently, one cannot say positively that the language explanation is the correct one. One way to shed further light on this issue might be to bridge the phylogenetic gap between monkeys and humans through tests on transitional species like the chimpanzee or bonobo. Apes would disconfirm the language hypothesis if they showed the human rank order of task difficulty. In turn, this raises the interesting possibility that the language hypothesis could be explored within the same species using ape participants who have and have not received language or symbolic training (Savage-Rumbaugh, 1986; Savage-Rumbaugh et al., 1993).

A related (possibly overarching) description of the species difference arises from current research in the cognitive neuroscience of category learning (Ashby, Alfonso-Reese, Turken, & Waldron, 1998; Ashby & Ell, 2001; F. G. Ashby, personal communication, January 2003). It is becoming increasingly clear that the cognitive function generally called categorization is served under different conditions (especially conditions of different category structure) by different brain regions. One important neural system for categorization in humans is an explicit system that is available to consciousness and that reasons deliberately about category organization and the solution to category problems. This system could serve the hypothesis-testing and rule-discovery processes that humans showed here and in Shepard et al. (1961) as they learned Type I and Type II tasks. This system is probably mediated by frontal cortical structures (e.g., prefrontal cortex and the anterior cingulate gyrus) that may serve more general functions of working memory and executive attention. These general functions would be well-suited to managing the processes of hypothesis memory, testing, switching, and so forth. Feldman's (2000) discussion of the logical incompressibility of the six task solutions may provide a clue to the general operating characteristics of this explicit system. It apparently appreciates Boolean simplicity.

The idea of this explicit or frontal category system maps well onto the human data pattern. Type I and Type II tasks would suit the rule-based operation of the explicit system, allowing this system to dominate learning in those cases and to produce rapid and sudden learning. In contrast, the explicit system might not support learning at all for Type VI tasks. It is interesting that Shepard et al. (1961) found that only Type VI performance in humans confirmed the predictions of conditioning and similarity-based theories (Figure 2A). Thus, one can also view Shepard et al.'s seminal study as one of the first surveys of category structures for the affordance they grant to explicit categorization strategies. Type I and Type II tasks seem to make explicit strategies highly feasible: Performance in these tasks falls far below the error-prediction line in Figure 2A. Type VI tasks do not, and Types III–V are intermediate (perhaps because they are learned through a combination of rules and exception–exemplar memorization).

This theoretical possibility fits with existing implicit–explicit dissociations within the categorization literature. For example, Smith et al. (1993) showed that depression differentially affected performance in different tasks. Performance in rule-based tasks (four-dimensional Type I tasks) was strongly impaired by depression. Performance in family-resemblance categorization tasks (four-dimensional Type IV tasks) was relatively preserved. Smith et al. concluded that depression had compromised humans' explicit category-learning system within which they would test hypotheses and find rules that solved the Type I task. (Smith et al.'s research on depression and other research in this area owes much to the original theoretical formulations made in Kemler Nelson, 1984, and Brooks, 1978). Waldron and Ashby (2001) also used tasks that essentially had a Type I and Type IV structure. They found that adding a concurrent cognitive load affected Type I task performance strongly, and Type IV task performance much less so. They also concluded that the Type I task draws heavily on an explicit category-learning system and that the concurrent task hijacks the cognitive resources that system requires.

Love (2002) made a similar finding regarding the Type II task. He gave human participants Shepard et al.'s (1961) tasks of Types I, II, IV, and VI under explicit and implicit category-learning conditions. Under explicit conditions, he found the usual rank order of task difficulty (I < II < IV < VI). However, under implicit conditions, he found essentially the rank order we found with monkeys here (I < IV < II < VI). Type II task performance became very difficult when the explicit categorization system was undercut.

This discussion potentially sheds a constructive light on the monkeys' performance and on the species difference between humans and monkeys. Monkeys are almost certainly weaker than humans regarding the cognitive capacities served by frontal brain

---

[3] For example, an economical rule for Type I would be if black, A; if white, B. For Type II, a parsimonious rule would be if black and triangle or white and square, A; if white and triangle or black and square, B. For Type VI, the rule would essentially be if large black triangle or small white triangle or large white square or small black square, A; if large white triangle or small black triangle or large black square or small white square, B. Note that this rule for Type VI is logically equivalent to describing each of the eight stimuli individually, supporting the idea that this task must be learned as an identification task in which labels are learned for eight noncategorizeable stimuli.

systems. Not only do monkeys have proportionally much smaller frontal cortices (Semendeferi, Lu, & Schenker, 2002), they are also compromised relative to humans on frontal tasks that offer response competition or require response inhibition (Stroop tasks, flanker tasks, etc.; A. C. Roberts, 1996; Washburn, 1994, 2003). In line with these results, it is possible that they have a rudimentary explicit category-learning system (F. G. Ashby, personal communication, January 2003).

This possibility is supported by the present findings. For example, on the basis of this analysis, monkeys should be especially impaired relative to humans on the Type II task because humans can solve it explicitly. That this is true is one of this article's central findings. For the same reason, monkeys should find the Type I task relatively harder than humans do. In fact, whereas Table 3 shows that humans overall had a 19.9% performance advantage for Type I over Type IV performance (94.6% vs. 74.7%), Table 4 shows that monkeys had only a 4.4% advantage (86.2% vs. 81.8%). Overall, if monkeys are substantially denied explicit category-learning processes, their scatter plots relating category structure to performance should take on the linear pattern that is consistent with a cognitive system that operates mainly by conditioning and stimulus-generalization processes. Figure 7 showed that their scatter plots do have this form.

We point out that there is probably productive neuroimaging research to be done on cross-species performance differences in categorization. The different tasks likely depend focally on different brain systems. The different species may use these brain systems in different ways for different tasks. In this domain, too, the Shepard et al. (1961) tasks offer a useful survey of tasks that encourage hypothesis testing and rule use and of tasks that will be learned using more implicit, generalization-based strategies. By also controlling for response modality, category size, and so forth, the Shepard et al. tasks are suitable for the subtractive designs favored in neuroimaging studies.

We also point out that this explicit–implicit explanatory framework stands in interesting juxtaposition to the language explanatory framework considered earlier. Presently, the ideas of explicit and verbally mediated category learning are treated almost synonymously because so much of the neuroscience research has used human participants. But important theoretical questions remain about whether explicit category learning is linked by convenience or by representational necessity to language in humans and about whether there are symbolic, semantic, or Boolean forms of explicit category learning that could be nonlinguistic in nature and that animals might use.

A final possible difference between monkeys and humans in categorization could lie in the robustness and agility of the metacognitive utility that oversees performance, monitors when performance is good, notices why performance is good, and emphasizes those attentional stances in performance. It is well-documented that humans have a cognitive monitor that oversees cognitive processing and problem solving and that they have a cognitive executive that directs information processing in more favorable directions (Dunlosky & Nelson, 1997; Koriat, 1993; Nelson, 1992, 1996; Schwartz, 1994). The study of cognitive regulation or metacognition in animals is a broadening field in comparative psychology (Hampton, 2001; Inman & Shettleworth, 1999; Shields, Smith, & Washburn, 1997; Smith, Schull, et al., 1995; Smith, Shields, Allendoerfer, & Washburn, 1998; Smith, Shields, Schull, & Wash-

burn, 1997; Smith, Shields, & Washburn, 2003a, 2003b). But it remains possible that animals are limited in this area and that they are slower to notice and to apply correct attentional strategies because of it. In fact, no research exists on animals' use of metacognitive monitoring in categorization tasks of the kind used here. This idea could be linked to the executive functions of attention already discussed from the perspective of neuroscience.

Regarding all the possible interpretations of the species difference, it is worth explaining why the species difference in the Type II task is so striking (cf. Figures 2 and 7). Somehow the Type II task is friendly to the human cognitive system but foreign to that of monkeys. The likely reason for this is that the Type II task has Boolean simplicity, making it friendly to verbalization, rule use, explicit solution efforts, and thus the human cognitive system. But perceptually the task is terrible. It is massively nonlinearly separable. It has a very complex decision boundary (Ashby & Gott, 1988; Ashby & Maddox, 1992). It defeats any attentional strategy that sums up category evidence across features to reach category decisions. It contains completely opposite stimuli (Figure 1) that nonetheless must be placed into the same category. It requires a highly contingent attentional strategy. That is, the polarity of attention must be shifted contingent on shape or on some other feature (e.g., if square, white = A and black = B; if triangle, black = A and white = B). Our guess is that these contingent shifts in the polarity of attention pose a difficult problem for an implicit categorization system that proceeds by cue–response associations or by generalization between exemplars. For all these reasons, the Type II task lacks perceptual coherence and lacks affordances for the processes of conditioning, and so it will be unfriendly to implicit solution efforts and thus to monkeys' cognitive system. The species difference should be striking because the Type II task maximally differentiates the feasibility of explicit and implicit processes.

We would not deny monkeys an explicit category-learning system altogether (for only one consideration, they have a prefrontal cortex). Indeed, it is possible that the present results do not do full justice to the extent to which monkeys could at the limit acquire rules suddenly, appreciate XOR relations, master contingent attentional strategies, and so forth. (Another possibility is that the present results may overestimate the extent to which humans normally do so in the world—see Brooks, 1978; Love, 2002; Kemler Nelson, 1984; Smith & Shapiro, 1989; Smith et al., 1993; Waldron & Ashby, 2001.) We have discussed the differences between the procedures used for humans and animals. These are cause to think carefully about the species difference so as to not interpret it too extremely. However, in our view, these differences are not cause to shy away from human–animal comparisons of the kind described here. Comparisons of this kind potentially embody some of the most important and intriguing questions about cognitive evolution; about the role of language, working consciousness, and executive functions in task performance; and about the brain regions that serve these functions.

As one route to evaluating monkeys' ultimate capacities in the rule-learning area, one could give them a series of consecutive category problems of the same type. In fact, this was part of what Shepard et al. (1961) did with humans. This research would give monkeys the chance to acquire learning sets for finding criterial attributes, learning sets for appreciating XOR

relations, and so forth. This research would thus bring to the domain of categorization the important ideas from the original learning-set studies (Harlow, 1949). In considering this possibility, it is worth remembering Harlow's theoretical motivation for studying learning sets. He thought that humans would arrive at the laboratory well-equipped with learning sets, having completed thousands of simple cognitive problems over the years, whereas the naive monkeys usually tested in cognitive studies would not be so equipped because they lacked this experience. By providing monkeys with serial experience in the different types of category task, one could ask how dominant a role one could finally get learning sets to play. This approach would be the converse of the ones already discussed that denied humans their learning sets by using some manipulation that produced less intentional or incidental learning.

However, despite the interest in alternative empirical approaches like this, it is noteworthy that all the monkeys, even across six rotations of six task types, instantiated so well Shepard et al.'s (1961) idea of a cognitive system that learns categories associatively through cue-conditioning processes or similarity-based generalization processes. Some within the behaviorist tradition will appreciate this outcome. Others within the cognitivist tradition will not. This outcome is not completely unexpected to us. Whereas our monkeys have sometimes shown symbolic or relational learning and sophisticated cognitive accomplishments (Shields et al., 1997; Smith et al., 1998, 2003a; Washburn, 1994; Washburn & Rumbaugh, 1991a), they have also sometimes shown equally dramatic failures of cognitive performance (Filion, 1993; Washburn, Gulledge, & Martin, 2003; Washburn, Smith, & Filion, 1998). Given the importance of categorization as a domain of current research and theory in cognitive psychology and its historical importance in theory and research in comparative psychology, it is a significant fact that monkeys seem in these tasks to learn more associatively than cognitively.

Thus, in a sense, the monkeys answer the important comparative question that Shepard et al. (1961) left unanswered 40 years ago: What cognitive system instantiates the associative or similarity-based categorization system they sought in humans but did not find? The present research suggests that some primate species may do so. This is an interesting point to consider from the standpoint of the evolutionary origins of humans' capacity for categorization, because it may reveal something about the stem categorization capacity from which that of humans emerged.

References

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.

Ashby, F. G., & Ell, S. W. (2001). The neurobiology of human category learning. *Trends in Cognitive Sciences, 5*, 204–210.

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. *Jour-*

nal of Experimental Psychology: Human Perception and Performance, 18, 50–71.

Astley, S. L., & Wasserman, E. A. (1992). Categorical discrimination and generalization in pigeons: All negative stimuli are not created equal. *Journal of Experimental Psychology: Animal Behavior Processes, 18*, 193–207.

Bhatt, R. S., Wasserman, E. A., Reynolds, W. F., & Knauss, K. S. (1988). Conceptual behavior in pigeons: Categorization of both familiar and novel examples from four classes of natural and artificial stimuli. *Journal of Experimental Psychology: Animal Behavior Processes, 14*, 219–234.

Brooks, L. R. (1978). Non-analytic concept formation and memory for instances. In E. Rosch & B. B. Lloyd (Eds.), *Cognition and categorization* (pp. 169–211). Hillsdale, NJ: Erlbaum.

Cerella, J. (1979). Visual classes and natural categories in the pigeon. *Journal of Experimental Psychology: Human Perception and Performance, 5*, 68–77.

Chase, S., & Heinemann, E. G. (2001). Exemplar memory and discrimination. In R. G. Cook (Ed.), *Avian visual cognition*. Retrieved January 2003 from http://www.pigeon.psy.tufts.edu/avc/chase/

Cheney, D. L., & Seyfarth, R. M. (1990). *How monkeys see the world*. Chicago: University of Chicago Press.

Cook, R. (Ed.). (2001). *Avian visual cognition*. Retrieved January 2003 from http://www.pigeon.psy.tufts.edu/avc/

D'Amato, M. R., & Van Sant, P. (1988). The person concept in monkeys (*Cebus apella*). *Journal of Experimental Psychology: Animal Behavior Processes, 14*, 43–55.

Dunlosky, J., & Nelson, T. O. (1997). Similarity between the cue for judgments of learning (JOL) and the cue for test is not the primary determinant of JOL accuracy. *Journal of Memory and Language, 36*, 34–39.

Feldman, J. (2000, October 5). Minimization of Boolean complexity in human concept learning. *Nature, 407*, 630–633.

Filion, C. M. (1993). *Extrapolation of trajectories by rhesus macaques*. Unpublished master's thesis, University of Georgia, Athens.

Flagg, S. F. (1974). Learning of the insoluble conditional reaction problem by rhesus monkeys. *Animal Learning and Behavior, 2*, 181–184.

Hampton, R. R. (2001, April 24). Rhesus monkeys know when they remember. *Proceedings of the National Academy of Sciences, 98*, 5359–5362.

Harlow, H. F. (1949). The formation of learning sets. *Psychological Review, 56*, 51–65.

Hayes, K. J. (1953). The backward curve: A method for the study of learning. *Psychological Review, 60*, 269–275.

Herrnstein, R. J., Loveland, D. H., & Cable, C. (1976). Natural concepts in pigeons. *Journal of Experimental Psychology: Animal Behavior Processes, 2*, 285–301.

Huber, L. (2001). Visual categorization in pigeons. In R. Cook (Ed.), *Avian visual cognition*. Retrieved January 2003 from http://www.pigeon.psy.tufts.edu/avc/huber/

Huber, L., & Lenz, R. (1993). A test of the linear feature model of polymorphous concept discrimination with pigeons. *Quarterly Journal of Experimental Psychology: Comparative and Physiological Psychology, 46B*, 1–18.

Inman, A., & Shettleworth, S. J. (1999). Detecting metamemory in nonverbal subjects: A test with pigeons. *Journal of Experimental Psychology: Animal Behavior Processes, 25*, 389–395.

Jitsumori, M. (1996). Prototype effect and categorization of artificial polymorphous stimuli in pigeons. *Journal of Experimental Psychology: Animal Behavior Processes, 22*, 405–418.

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

Koriat, A. (1993). How do we know that we know? The accessibility model of the feeling of knowing. *Psychological Review*, *100*, 609–639.

Krechevsky, L. (1932). "Hypotheses" in rats. *Psychological Review*, *39*, 516–532.

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

Lea, S. E. G., & Ryan, C. M. E. (1990). Unnatural concepts and the theory of concept discrimination in birds. In M. L. Commons, R. J. Herrnstein, S. M. Kosslyn, & D. B. Mumford (Eds.), *Quantitative analyses of behavior* (Vol. 8, pp. 165–185). Hillsdale, NJ: Erlbaum.

Love, B. C. (2002). Comparing supervised and unsupervised category learning. *Psychonomic Bulletin & Review*, *9*, 829–835.

Love, B. C., & Medin, D. L. (1998). SUSTAIN: A model of human category learning. In *Proceedings of the Fifteenth National Conference on Artificial Intelligence and Tenth Conference on Innovative Applications of Artificial Intelligence* (pp. 671–676). Menlo Park, CA: AAAI Press.

Medin, D. L. (1975). A theory of context in discrimination learning. In G. H. Bower (Ed.), *The psychology of learning and motivation* (Vol. 9, pp. 269–315). New York: Academic Press.

Medin, D. L., Altom, M. W., & Murphy, T. D. (1984). Given versus induced category representations: Use of prototype and exemplar information in classification. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *10*, 333–352.

Medin, D. L., & Dewey, G. I. (1984). Learning of ill-defined categories by monkeys. *Canadian Journal of Psychology*, *38*, 285–303.

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

Nelson, T. O. (Ed.). (1992). *Metacognition: Core readings*. Toronto, Ontario, Canada: Allyn & Bacon.

Nelson, T. O. (1996). Consciousness and metacognition. *American Psychologist*, *51*, 102–116.

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

Nosofsky, R. M., Gluck, M. A., Palmeri, T. J., McKinley, S. C., & Glauthier, P. (1994). Comparing models of rule-based classification learning: A replication and extension of Shepard, Hovland, and Jenkins (1961). *Memory & Cognition*, *22*, 352–369.

Pearce, J. M. (1987). A model of stimulus generalization for Pavlovian conditioning. *Psychological Review*, *94*, 61–73.

Pearce, J. M. (1988). Stimulus generalization and the acquisition of categories by pigeons. In L. Weiskrantz (Ed.), *Thought without language* (pp. 132–152). Oxford, England: Oxford University Press.

Pearce, J. M. (1994a). Discrimination and categorization. In N. J. Mackintosh (Ed.), *Animal learning and cognition* (pp. 109–134). New York: Academic Press.

Pearce, J. M. (1994b). Similarity and discrimination: A selective review and a connectionist model. *Psychological Review*, *101*, 587–607.

Roberts, A. C. (1996). Comparison of cognitive function in human and nonhuman primates. *Cognitive Brain Research*, *3*, 319–327.

Roberts, W. A., & Mazmanian, D. S. (1988). Concept learning at different levels of abstraction by pigeons, monkeys, and people. *Journal of Experimental Psychology: Animal Behavior Processes*, *14*, 247–260.

Rumbaugh, D. M., & Pate, J. L. (1984). The evolution of cognition in primates: A comparative perspective. In H. L. Roitblat, T. G. Bever, & H. S. Terrace (Eds.), *Animal cognition* (pp. 569–589). Hillsdale, NJ: Erlbaum.

Rumbaugh, D. M., Richardson, W. K., Washburn, D. A., Savage-Rumbaugh, E. S., & Hopkins, W. D. (1989). Rhesus monkeys (*Macaca mulatta*), video tasks, and implications for stimulus–response spatial contiguity. *Journal of Comparative Psychology*, *103*, 32–38.

Savage-Rumbaugh, E. S. (1986). *Ape language: From conditioned response to symbol*. New York: Columbia University Press.

Savage-Rumbaugh, E. S., Murphy, J., Sevcik, R. A., Brakke, K. E., Williams, S. L., & Rumbaugh, D. M. (1993). Language comprehension in ape and child. *Monographs of the Society for Research in Child Development*, *58*(3–4, Serial No. 233).

Schrier, A. M., & Brady, P. M. (1987). Categorization of natural stimuli by monkeys (*Macaca mulatta*): Effects of stimulus set size and modification of exemplars. *Journal of Experimental Psychology: Animal Behavior Processes*, *13*, 136–143.

Schwartz, B. L. (1994). Sources of information in metamemory: Judgments of learning and feelings of knowing. *Psychonomic Bulletin & Review*, *1*, 357–375.

Semendeferi, K., Lu, A., & Schenker, N. (2002). Humans and great apes share a large frontal cortex. *Nature Neuroscience*, *5*, 272–276.

Shepard, R. N., Hovland, C. I., & Jenkins, H. M. (1961). Learning and memorization of classifications. *Psychological Monographs*, *75*(13, Whole No. 517).

Shields, W. E., Smith, J. D., & Washburn, D. A. (1997). Uncertain responses by humans and rhesus monkeys (*Macaca mulatta*) in a psychophysical same–different task. *Journal of Experimental Psychology: General*, *126*, 147–164.

Smith, J. D., Schull, J., Strote, J., McGee, K., Egnor, R., & Erb, L. (1995). The uncertain response in the bottlenosed dolphin (*Tursiops truncatus*). *Journal of Experimental Psychology: General*, *124*, 391–408.

Smith, J. D., & Shapiro, J. H. (1989). The occurrence of holistic categorization. *Journal of Memory and Language*, *28*, 386–399.

Smith, J. D., Shields, W. E., Allendoerfer, K. R., & Washburn, W. A. (1998). Memory monitoring by animals and humans. *Journal of Experimental Psychology: General*, *127*, 227–250.

Smith, J. D., Shields, W. E., Schull, J., & Washburn, D. A. (1997). The uncertain response in humans and animals. *Cognition*, *62*, 75–97.

Smith, J. D., Shields, W. E., & Washburn, D. A. (2003a). The comparative psychology of uncertainty monitoring and metacognition. *Behavioral and Brain Sciences*, *26*, 317–339.

Smith, J. D., Shields, W. E., & Washburn, D. A. (2003b). Inaugurating a new area of comparative cognition research. *Behavioral and Brain Sciences*, *26*, 358–373.

Smith, J. D., Tracy, J., & Murray, M. J. (1993). Depression and categorization. *Journal of Experimental Psychology: General*, *122*, 331–346.

Struhsaker, T. T. (1967). Auditory communication among vervet monkeys (*Cercopithecus aethiops*). In S. A. Altmann (Ed.), *Social communication among primates* (pp. 281–324). Chicago: University of Chicago Press.

Trabasso, T., & Bower, G. H. (1968). *Attention in learning: Theory and research*. New York: Wiley.

Vaughn, W., Jr., & Greene, S. L. (1984). Pigeon visual memory capacity. *Journal of Experimental Psychology: Animal Behavior Processes*, *10*, 256–271.

von Fersen, L., & Lea, S. E. G. (1990). Category discrimination by pigeons using five polymorphous features. *Journal of the Experimental Analysis of Behavior*, *54*, 69–84.

Waldron, E. M., & Ashby, F. G. (2001). The effects of concurrent task interference on category learning: Evidence for multiple category systems. *Psychonomic Bulletin & Review*, *8*, 168–176.

Washburn, D. A. (1994). Stroop-like effects for monkeys and humans: Processing speed or strength of association? *Psychological Science*, *5*, 375–379.

Washburn, D. A. (2003). *Executive versus experiential constraints on attention*. Manuscript submitted for publication.

Washburn, D. A., Gulledge, J. P., & Martin, B. A. (2003). A species difference in visuospatial memory: A failure of memory for what, where, or what is where? *International Journal of Comparative Psychology*, *16*, 209–225.

Washburn, D. A., & Rumbaugh, D. M. (1991a). Ordinal judgments of Arabic symbols by macaques (*Macaca mulatta*). *Psychological Science, 2*, 190–193.

Washburn, D. A., & Rumbaugh, D. M. (1991b). Testing primates with joystick-based automated apparatus: Lessons from the Language Research Center's computerized test system. *Behavior Research Methods, Instruments, & Computers, 24*, 157–164.

Washburn, D. A., Smith, J. D., & Filion, C. M. (1998, August). *"Meaningful failures" in learning*. Presented at the 106th Annual Convention of the American Psychological Association, San Francisco, CA.

Wasserman, E. A., Kiedinger, R. E., & Bhatt, R. S. (1988). Conceptual behavior in pigeons: Categories, subcategories, and pseudocategories. *Journal of Experimental Psychology: Animal Behavior Processes, 14*, 235–246.

Zeaman, D., & House, B. J. (1963). The role of attention in retardate discrimination learning. In N. R. Ellis (Ed.), *Handbook of mental deficiency* (pp. 159–223). New York: McGraw-Hill.

New Editors Appointed, 2006–2011

The Publications and Communications Board of the American Psychological Association announces the appointment of seven new editors for 6-year terms beginning in 2006. As of January 1, 2005, manuscripts should be directed as follows:

- *Experimental and Clinical Psychopharmacology* (www.apa.org/journals/pha.html), **Nancy K. Mello, PhD,** McLean Hospital, Massachusetts General Hospital, Harvard Medical School, 115 Mill Street, Belmont, MA 02478-9106.

- *Journal of Abnormal Psychology* (www.apa.org/journals/abn.html), **David Watson, PhD,** Department of Psychology, University of Iowa, Iowa City, IA 52242-1407.

- *Journal of Comparative Psychology* (www.apa.org/journals/com.html), **Gordon M. Burghardt, PhD,** Department of Psychology or Department of Ecology & Evolutionary Biology, University of Tennessee, Knoxville, TN 37996.

- *Journal of Counseling Psychology* (www.apa.org/journals/cou.html), **Brent S. Mallinckrodt, PhD,** Department of Educational, School, and Counseling Psychology, 16 Hill Hall, University of Missouri, Columbia, MO 65211.

- *Journal of Experimental Psychology: Human Perception and Performance* (www.apa.org/journals/xhp.html), **Glyn W. Humphreys, PhD,** Behavioural Brain Sciences Centre, School of Psychology, University of Birmingham, Edgbaston, Birmingham B15 2TT, United Kingdom.

- *Joural of Personality and Social Psychology: Attitudes and Social Cognition* section (www.apa.org/journals/psp.html), **Charles M. Judd, PhD,** Department of Psychology, University of Colorado, Boulder, CO 80309-0345.

- *Rehabilitation Psychology* (www.apa.org/journals/rep.html), **Timothy R. Elliott, PhD,** Department of Psychology, 415 Campbell Hall, 1300 University Boulevard, University of Alabama, Birmingham, AL 35294-1170.

Electronic submission: As of January 1, 2005, authors are expected to submit manuscripts electronically through the journal's Manuscript Submission Portal (see the Web site listed above with each journal title).

Manuscript submission patterns make the precise date of completion of the 2005 volumes uncertain. Current editors, Warren K. Bickel, PhD, Timothy B. Baker, PhD, Meredith J. West, PhD, Jo-Ida C. Hansen, PhD, David A. Rosenbaum, PhD, Patricia G. Devine, PhD, and Bruce Caplan, PhD, respectively, will receive and consider manuscripts through December 31, 2004. Should 2005 volumes be completed before that date, manuscripts will be redirected to the new editors for consideration in 2006 volumes.