JOURNAL OF THE EXPERIMENTAL ANALYSIS OF BEHAVIOR 2002, **78,** 333–343 NUMBER 3 (NOVEMBER)

FAILURE TO FIND EVIDENCE OF STIMULUS GENERALIZATION WITHIN PICTORIAL CATEGORIES IN PIGEONS

JENNIFER E. SUTTON AND WILLIAM A. ROBERTS

UNIVERSITY OF WESTERN ONTARIO

Pigeons' key pecks were reinforced in the presence of pictures from one of two categories, cats or cars. A single picture associated with reinforcement was used in Experiment 1, and 20 pictures from the same category were associated with reinforcement in Experiment 2. Pigeons then were presented with novel test pictures from the training category and from the other, previously unseen, category. During Session 1 of testing, pigeons pecked no more often at pictures from the reinforced category than at pictures from the previously unseen category. When pigeons were trained with pictures associated with reinforcement or its absence from different categories in Experiment 3, differential responding to novel pictures from different categories appeared during Session 1. These findings argue against a process of automatic stimulus generalization within natural categories and in favor of the position that category distinctions are not made until members of at least two categories are compared with one another.

Key words: concepts, categories, stimulus generalization, discrimination, pictures, pigeons

Studies showing that pigeons and other animals readily acquire concepts based on response to pictures of natural and artificial objects have received considerable attention. In groundbreaking work, Herrnstein and Loveland (1964) reinforced key pecking of pigeons whenever a picture containing people was shown, and pecking the key was not reinforced whenever a picture without people was shown. When novel pictures with people present or absent were presented, pigeons continued to peck more at people pictures than at nonpeople pictures. Later studies showed that pigeons responded differentially to pictures that did and did not contain trees, bodies of water, and images of a particular person (Herrnstein, Loveland, & Cable, 1976) or to underwater pictures that did or did not contain fish (Herrnstein & de Villiers, 1980).

Herrnstein interpreted these studies to suggest that pigeons saw two-dimensional pictures presented on a screen as three-dimensional objects and that they saw these objects as falling into classes similar to those seen by people and described by verbal categories. Thus, he argued that

> Our data support at least a partial isomorphy in the inferred classes between pigeons and people. (Herrnstein et al., 1976, p. 301)

> It would be hard to argue that the subjects formed the relevant categories here by means of differential reinforcement or any other correlate of presentation time. If the categories were not formed here, they must have been preexisting. . . . Their performance suggests that they are seeing the stimuli approximately as we do, as representing a three-dimensional space containing solid objects. (Herrnstein, 1979, p. 125)

> Something in the pigeon's perceptual dynamics ties fish together as a class, prior to differential reinforcement. (Herrstein & de Villiers, 1980, p. 87)

> What the experimenter had in mind converged fairly well with what the pigeon evidently used as a principle of categorizing. When there is such a convergence, categorization should speed discrimination. (Herrnstein, 1984, p. 236)

Categorization was assumed to occur immediately upon perception of pictures and prior to behavioral discrimination promoted by differential reinforcement.

In a novel procedure that extended these studies, Bhatt, Wasserman, Reynolds, and Knauss (1988) showed pigeons images of pictures from four different categories: cats, cars, flowers, and chairs. A key was placed at

Support for the research reported in this article was provided by a research grant from the Natural Sciences and Engineering Research Council of Canada.

Correspondence concerning this article should be addressed to William A. Roberts, Department of Psychology, University of Western Ontario, London, Ontario N6A 5C2, Canada (e-mail: roberts@uwo.ca).

each corner of the screen on which pictures were displayed, and each category required a peck on a different key for reinforcement. When the same set of pictures was shown on repeated sessions, pigeons learned to peck the correct key at about 80% accuracy, compared with a chance level of 25%. Even more impressive than this finding was the demonstration that the pigeons learned to peck the correct key almost as rapidly when the set of pictures from each category was novel on each session. Thus, pigeons were not learning to make correct responses to a fixed set of stimuli; responses to category members generalized to never-before-seen pictures. When *pseudocategories* were formed by requiring different responses for reinforcement to categories formed by combining pictures of cats, cars, flowers, and chairs, pigeons learned far more slowly than they did when each category contained only pictures of cats, cars, flowers, or chairs (Wasserman, Kiedinger, & Bhatt, 1988; see also Herrnstein & de Villiers, 1980). When correct responses to two subsets of pictures from a category such as cats required responses to different keys, pigeons often made *conceptual errors* by placing one cat picture in the wrong category containing other cat pictures (Wasserman et al., 1988). Discrimination among categories of pictures recognized as natural concepts by people and given different verbal labels by them seemed to be particularly easy for pigeons to acquire (Herrnstein et al., 1976; Wasserman, 1995).

Reinforcement of pictures within a category also can retard the development of inhibitory responding to other pictures in the same category. Astley and Wasserman (1992) presented pigeons with 12 pictures from each of four categories (people, flowers, cars, and chairs). Responses to all 48 pictures were nonreinforced, whereas responses to 12 additional pictures from only one of these categories were reinforced. Pigeons learned to peck at reinforced pictures and not to peck at nonreinforced pictures, but pigeons learned more slowly not to peck at the nonreinforced pictures that fell into the same category as the reinforced pictures. Based on these findings and the tendency of pigeons to make conceptual errors when discriminating pseudocategories, they concluded that "the data are consistent with the proposal of Herrnstein and de Villiers (1980) that perceptual mechanisms are primarily responsible for the pigeon's tendency to treat members of one conceptual class differently from members of other conceptual classes" (Astley & Wasserman, p. 202). Astley and Wasserman further developed these ideas into a general account of categorical discrimination and transfer. Their account borrowed heavily from Spence's (1936, 1937) theory of discrimination learning. Spence held that reinforcement delivered for a response to a stimulus value at a particular point on a dimension established a tendency for that stimulus to excite that response on subsequent presentations. Spence also held, as did Pavlov (1927), that excitation spread to other points along the stimulus dimension, with progressive weakening at increasing distances from the reinforced value. Spence's theory has been applied largely to dimensions that vary in a single physical property, such as wavelength of light, frequency of a tone, or angular tilt of a straight line. Astley and Wasserman recognized that the pictures used in natural concept experiments were too complex to be scaled along a single dimension. Pictures of natural objects vary along many dimensions and may be hypothesized to be represented in an *n*-dimensional space, representing all the values on all the dimensions contained in the pictures. Nevertheless, using a distance metric of similarity, it was assumed that members of the same category would lie closer to one another within such a semantic space than they would to the members of a different category. Thus, stimulus generalization of responses would be greater among pictures that look similar to one another than among pictures that do not look similar to one another.

Given this fundamental assumption about primary stimulus generalization within and among pictures of objects from natural categories, Astley and Wasserman (1992) proposed that picture stimuli acquired excitatory and inhibitory control of responses through repeated reinforcement and nonreinforcement. Both excitation and inhibition generalized strongly to similar pictures (those within the same category as the target stimulus) but generalized only weakly to dissimilar pictures (those in other categories). These assumptions account for a number of findings from natural concept studies (Roberts, 1998).

They explain why animals trained to discriminate among sets of stimuli from different categories can accurately classify novel pictures from those categories and why discrimination takes place faster among natural categories than among pseudocategories. The theory also explains why responses to novel stimuli increase in accuracy as the number of training stimuli within categories increases (Wasserman & Bhatt, 1992). As more stimuli are trained, the likelihood that a novel exemplar will resemble a training stimulus, and thus benefit from excitatory generalization from that stimulus, increases.

Although both Herrnstein and Wasserman and their colleagues argued that pigeons immediately perceived pictures within categories as more similar to one another than pictures in different categories, we argue that no definitive test of this hypothesis has been made. Pigeons' ability to sort pictures into categories, their commission of conceptual errors, and their increased errors with pseudocategories and nonreinforced pictures that fall into the same category as reinforced pictures show more generalization within than among picture categories. These effects appeared, however, after pigeons had begun discrimination training with reinforced and nonreinforced responses to different pictures. It may be argued that the process of differential reinforcement causes pigeons to attend to differences between pictures and thus to learn quickly about the features that define categories. The question we address is whether pigeons perceive categorical organization of pictures before differential reinforcement occurs.

One previous experiment reported by Cerella (1979, Experiment 4) suggests that pigeons might not recognize a category prior to differential reinforcement, but its interpretation is ambiguous. Key pecking of pigeons was reinforced when they were shown repeated pictures of the same white-oak leaf for 40 trials. After nine sessions, they were tested with the 40 training pictures, 20 pictures of new white-oak leaves, and 20 pictures of non-oak leaves (pictures of leaves from other species of trees). The mean pecks per minute to positives, other oak leaves, and non-oak leaves were 49, 54, and 46, respectively, and did not differ significantly. Cerella's interpretation of these findings was that pigeons had not attended to the single oak-leaf images during training and thus had learned nothing that could form the basis for discrimination between the test pictures. He then modified the apparatus so that pigeons had to peck directly on the screen containing the picture. The *same* pigeons used in Phase 1 now were given eight more training sessions, each containing 40 trials of reinforced pecking on the single positive white-oak leaf followed by another test session. On Test 2, the mean pecks per minute to positives, other white-oak leaves, and non-oak leaves were 51, 48, and 8, respectively, and showed significant discrimination between white-oak leaves and non-oak leaves. Although the difference in findings between Test 1 and Test 2 can be explained by assuming that pigeons attended to pictures in Phase 2 training but not in Phase 1 training, these findings also may be explained by the assumption that pigeons did not perceive white-oak leaves as a category until pecking was differentially reinforced for white-oak leaves and non-oak leaves. Although pigeons may have attended to the repeated reinforced presentations of the single white-oak leaf picture in Phase 1, this training may not have promoted any discrimination between white-oak leaves and non-oak leaves. During Test 1, however, pigeons were exposed to reinforced presentations of the training white-oak leaf and nonreinforced presentations of 20 novel white-oak leaves and 20 novel non-oak leaves. The pigeons, now having experienced differentially reinforced images from different leaf categories, showed discrimination between the reinforced category and the nonreinforced category on Test 2.

The experiments reported here used pictures from two categories, cats and cars. A pigeon's key pecks were reinforced in the presence of pictures in one of these categories, and then the pigeon was tested for rate of response to novel pictures from the reinforced category and novel pictures from the unseen category. Of importance, pigeons pecked directly on a touch screen containing the picture shown, thus ensuring attention to the picture. The theory of stimulus generalization makes an important prediction about this experiment. If reinforcement of responses to pictures in one category establishes a tendency to respond to those pictures that automatically generalizes to similar depictions of members of that category, pigeons should respond more strongly to novel pictures from the reinforced category than to novel pictures from the unseen category. On the other hand, if categories are not formed until pigeons' pecks are differentially reinforced in the presence of pictures in different categories, they should not respond differentially to stimuli in the same or different categories.

EXPERIMENT 1

In Experiment 1, pigeons' pecks were reinforced in daily sessions in the presence of a single picture. For half the pigeons, the picture was that of a cat, and for the other half, the picture was that of a car. After pigeons had learned to peck at these pictures at a high rate, each subgroup was tested for response to 10 novel cat pictures and 10 novel car pictures. The theory of stimulus generalization predicts that pigeons receiving food for pecking at a cat picture should respond more often to novel cat pictures than to novel car pictures and that pecking associated with food delivery in the presence of a car picture should occur more often in the presence of novel car pictures than in the presence of novel cat pictures.

Method

Pigeons. Six adult White Carneau pigeons (*Columba livia*) served as subjects. The pigeons were hatched and raised in captivity; they had varying degrees of experience in operant chamber tasks on time processing, and some had participated in an open-arena task on landmark learning. They were maintained at approximately 85% of their free-feeding body weights during the experiment through mixed grain obtained in daily test sessions and supplemental pigeon chow in their home cages. Each pigeon was housed in a wire mesh cage with constant access to water and health grit. Overhead fluorescent lights in the pigeon colony room were illuminated from 7:00 a.m. to 7:00 p.m. each day, and pigeons were tested 5 days per week between 9:00 a.m. and 4:00 p.m.

Apparatus. The experimental chamber was constructed of Plexiglas and measured 46 cm by 38 cm by 36 cm. On the front wall, an opening (28 cm by 20 cm) allowed access to a color monitor (Zenith ZCM-1492, St. Joseph, MI) fitted with an infrared touch frame (Carroll Touch 1492 Smart Frame, Round Rock, TX). Between the frame and the surface of the monitor was a Plexiglas sheet that shielded the monitor from direct contact during testing. Access to mixed grain was provided through an aperture (1.5 cm diameter) located on the floor, 3 cm from the front wall and 16 cm from both the left and right walls. A small lamp under the floor illuminated the aperture as the feeder was presented. Presentation of graphics and operation of the feeder were controlled from a Pentium® PC and a locally constructed interface located in an adjacent room. Two different training stimuli were used. Three pigeons were trained with an image of a cat, and 3 were trained with an image of a car. Both images were in color and measured 435 by 280 pixels (about 18 cm by 11.5 cm) when presented on the monitor.

Procedure. Pigeons were pretrained to eat from the hopper by leaving the hopper raised with the overhead houselight illuminated for one or two daily sessions, or until the pigeon ate from the hopper. Each pigeon then was autoshaped to peck a white rectangle stimulus (18 cm by 11.5 cm) presented on a dark gray background on the screen. The stimulus was presented for 8 s or until the pigeon pecked within the stimulus boundary, whichever came first. The screen then went blank and the hopper was raised for 5 s, followed by a 10-s intertrial interval (ITI). The houselight was not illuminated at any point in the trial. Autoshaping sessions consisted of 90 trials. After a pigeon pecked reliably in autoshaping sessions, a fixed-ratio (FR) requirement was introduced and was gradually increased until the pigeon was required to peck the rectangle 10 times on each trial for reinforcement. In addition, the hopper access time was shortened to 3 s. Once pigeons had completed a session of 80 trials at FR 10, the formal training phase commenced.

Training sessions consisted of 90 trials. On each trial, the white rectangle was presented initially. After the pigeon pecked the white rectangle once, it was turned off and the training stimulus was presented (a cat for 3 pigeons, a car for the other 3 pigeons). The first peck after 15 s on the image turned off the stimulus and was followed by 3 s of reinforcement and then a 10-s ITI, during which

Table 1

Peck rates (pecks per minute) on five test sessions to the S+ picture, to novel pictures from the same category as S+, and to novel pictures from the unseen category (Experiment 1).

	Test pictures														
	Training S+					Same category as S+					Unseen category				
Pigeon	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5
8082 (car)	12.1	136.3	119.9	107.0	122.8	34.0	12.4	8.8	10.8	10.4	34.8	18.4	9.6	13.2	9.6
9936 (car)	93.0	79.6	82.3	84.3	84.8	20.0	6.0	1.6	2.4	2.0	18.4	7.6	6.4	8.8	7.6
M20 (car)	52.3	47.1	55.4	53.1	62.7	8.4	1.2	0.4	0.8	2.4	7.6	7.2	1.6	2.0	5.6
M2 (cat)	31.2	28.7	30.6	34.6	33.5	38.4	21.2	18.4	16.8	10.4	25.6	10.8	9.2	6.4	6.4
M1 (cat)	55.4	62.3	66.5	54.8	61.4	48.8	11.6	13.2	6.4	12.0	43.6	9.6	9.2	2.0	3.6
4820 (cat)	72.7	76.6	64.7	66.6	67.0	50.0	37.2	28.8	14.0	13.2	28.0	16.8	20.8	2.8	6.0
M	70.9	71.7	69.9	66.7	72.0	33.3	14.9	11.9	8.5	8.4	26.3	11.7	9.5	5.9	6.5
SEM	12.0	13.7	11.1	9.6	11.1	6.1	4.8	4.0	2.4	1.8	4.7	1.8	2.4	1.7	0.8

the monitor screen was dark gray and the chamber was dark. At the start of training, pigeons' pecks were reinforced on 100% of the trials in a training session. Over the training period, the probability of reinforcement was gradually reduced to 70% on each trial. Pigeons received one training session per day for 10 days.

Test sessions consisted of 100 trials. On one trial in every block of five trials, a novel test stimulus was presented, and the training stimulus was presented on the other four trials; responses to the training stimulus continued to be reinforced on 70% of the trials. Each

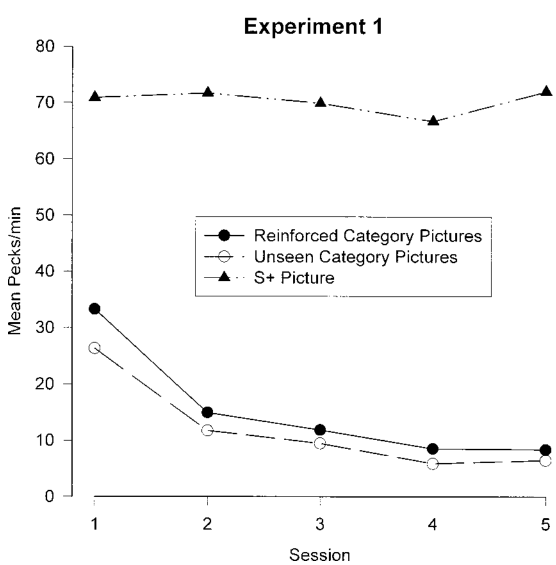

Fig. 1. Peck rates on novel test pictures from the reinforced training category and from the previously unseen category plotted over five sessions (Experiment 1).

novel stimulus was presented for 15 s and was then turned off without reinforcement whether the pigeon pecked or not. The position of the novel trial within the block of training trials was random across blocks. The novel test stimuli were 10 cat pictures and 10 car pictures that were presented in color and were the same size as the training stimulus. Each test stimulus was presented once in a test session. All pigeons were tested with the same 20 novel stimuli. The novel stimuli were presented in a different random order each session. The pigeons were tested for five sessions, during which the numbers of pecks to all stimuli were recorded. Throughout these experiments, effects were considered significant when $p < .05$.

Results and Discussion

The peck rates rose over the 10 sessions of training. In the group shown the cat picture, the mean peck rates during Sessions 1 and 10 were 41.8 and 63.1, respectively. In the group shown the car picture, the mean peck rates during Sessions 1 and 10 were 62.1 and 80.6, respectively. An analysis of variance (ANOVA), however, yielded nonsignificant effects of session, $F(9, 36) = 2.13$, and of group and the Group × Session interaction, $F\text{s} < 1.00$.

Table 1 shows the rates of pecking for each pigeon to test pictures over five test sessions, and Figure 1 shows the mean rates of pecking to the test pictures over the five test sessions. The reinforced-category curve shows rate of pecking at novel pictures in the same category as the training-reinforced picture (cat pictures for cat-reinforced pigeons and car pictures for car-reinforced pigeons), and the unseen-category curve shows rate of pecking at novel pictures in the category different from the training-reinforced picture (cat pictures for car-reinforced pigeons and car pictures for cat-reinforced pigeons). Rate of response to the training positive stimulus (S+) was substantially higher than to the novel pictures, and responses to the novel pictures from the reinforced category appeared to differ little from responses to pictures from the unseen category.

A one-way ANOVA was performed to compare the three picture conditions (S+ picture, pictures from the reinforced category, and pictures from the unseen category) on Test Session 1, and it yielded a significant effect of stimulus condition, $F(2, 10) = 8.26$. Newman–Keuls tests used to compare pairs of means showed that response to the S+ stimulus was significantly higher than to the novel pictures from both of the novel pictures categories and that the response rates to novel pictures from the reinforced category and the unseen category did not differ significantly. An ANOVA revealed significant effects of picture condition, $F(2, 10) = 19.54$, session, $F(4, 20) = 12.98$, and the Picture Condition × Session interaction, $F(8, 40) = 5.93$. The session and interaction effects arose from the drop in responding over sessions to the novel pictures in the reinforced and unseen categories while rate of response to the S+ picture remained consistently high. Newman–Keuls tests showed that pigeons pecked the S+ picture at a significantly higher rate than pictures in either of the novel categories, and rates of response to novel pictures in the reinforced and unseen categories did not differ significantly.

These results fail to yield any evidence of stimulus generalization within the cat or car categories. Reinforcement of pecking at a picture in one category for 10 sessions did not result in pecking significantly more frequently at novel pictures in the same category than they did at novel pictures in a category never before seen.

EXPERIMENT 2

A concern with the procedure used in Experiment 1 is that pigeons were trained to peck at only one picture from the cat or car category. If this training established a tendency to respond only to that picture and others similar to it within a semantic space, many of the test pictures within the training category may not have been sufficiently similar to the training picture to benefit from stimulus generalization. In support of this concern, Wasserman and Bhatt (1992) found that pigeons' ability to respond correctly to new pictures from four categories increased from the near-chance level of 27% correct when training involved one picture per category to about 65% correct when training involved 12 pictures per category. Thus, the failure to find differential generalization in Experiment 1 may have arisen from the use of only one training picture and not from the absence of stimulus generalization.

In Experiment 2, this possibility was examined by training pigeons with multiple exemplars from each of the training categories used in Experiment 1. Pecking by pigeons in one group was reinforced in the presence of 20 different pictures of cats, and for another group of pigeons, pecking was reinforced in the presence of 20 different pictures of cars. After 20 training sessions, each group was tested with 10 novel pictures from each category randomly interspersed among the training pictures.

Method

Pigeons. Six different adult pigeons served as subjects. Their research histories were similar to the pigeons in Experiment 1. They were housed and maintained in a manner identical to that in Experiment 1.

Apparatus. The touch-screen apparatus was the same as in Experiment 1. The training images were 20 different cats (for 3 pigeons) and 20 different cars (for the other 3 pigeons). The images were the same size as those used in Experiment 1. The test stimuli were the same 20 novel test images used in Experiment 1.

Procedure. The pigeons were pretrained as in Experiment 1. Training sessions consisted of 100 trials. In each session, 20 different cat images (for 3 pigeons) or 20 different car images (for the other 3 pigeons) were presented five times. The order of image presentation was random. The stimuli were presented for 15 s, and the first peck after 15 s resulted in reinforcement. The probability of reinforce-

Table 2

Peck rates (pecks per minute) on five test sessions to the S+ pictures, to novel pictures from the same category as the S+, and to novel pictures from the unseen category (Experiment 2).

	Test pictures														
	Training S+					Same category as S+					Unseen category				
Pigeon	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5
8 (car)	33.2	30.8	27.1	38.8	44.6	35.2	32.0	26.4	35.6	48.0	29.6	31.6	20.0	30.8	27.2
10326 (car)	41.3	41.0	39.2	28.0	34.8	31.6	30.4	38.0	29.6	25.6	47.2	45.2	35.2	18.4	8.4
4809 (car)	72.0	78.3	82.6	70.2	62.7	56.4	85.6	90.8	72.8	60.8	88.4	92.0	58.8	84.8	66.8
18193 (cat)	52.3	39.4	49.8	51.2	40.1	46.0	31.6	30.0	38.0	34.0	56.8	55.6	19.6	33.6	26.4
4807 (cat)	76.6	89.5	86.9	90.6	97.4	80.4	80.0	95.6	72.0	84.8	74.0	48.0	57.2	34.4	45.2
3869 (cat)	59.8	42.1	64.6	63.8	60.7	63.6	34.4	44.0	32.0	38.0	45.2	19.2	28.0	15.6	16.0
M	55.9	53.5	58.4	57.1	56.7	52.2	49.0	54.1	46.7	48.5	56.9	48.6	36.5	36.3	31.7
SEM	6.3	9.0	8.9	8.4	8.5	6.9	9.8	11.5	7.5	8.0	7.9	9.3	6.6	9.4	7.9

ment was gradually dropped from 100% at the beginning of training to 70% at the end of the training phase. Pigeons received 20 training sessions, with one session conducted each day.

Test sessions also consisted of 100 trials. As in Experiment 1, one trial in each block of five was a novel-image trial, and the other four trials were images the pigeons had seen in training. Responding in the presence of the training images was reinforced 70% of the time; the novel pictures were each presented for 15 s, and their presentation was never followed by reinforcement. The novel images were the same pictures used in Experiment 1, and all pigeons were tested with all 20 novel pictures (10 were cats and 10 were cars). The position of the novel picture in each block was random, and the order of the test pictures over the session was random. Pigeons received five test sessions.

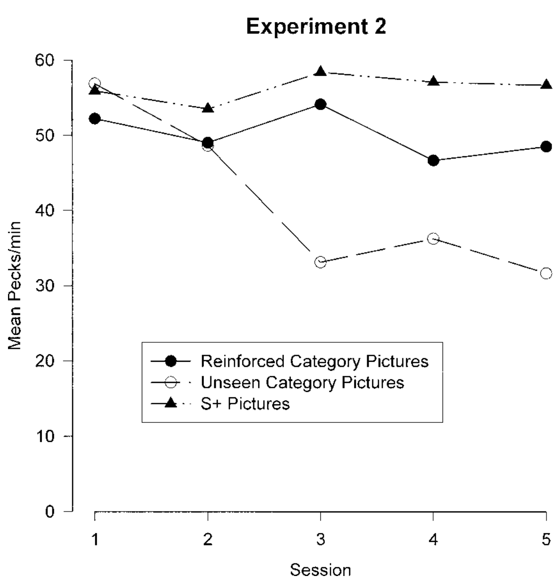

Fig. 2. Peck rates on novel test pictures from the reinforced training category and from the previously unseen category plotted over five sessions (Experiment 2).

Results and Discussion

The mean peck rates rose over the 20 sessions of training. The mean peck rates of the pigeons in the group in which cats were the S+ were 40.8 during Session 1 and 68.3 during Session 20. The mean peck rates of the pigeons in the group in which cars were the S+ were 33.9 during Session 1 and 44.6 during Session 20. An ANOVA, however, revealed no significant effects of group, $F < 1.00$, session, $F(19, 76) = 1.21$, or the Group × Session interaction, $F(19, 76) = 1.48$.

Table 2 shows peck rates for each pigeon over five test sessions, and Figure 2 shows the mean peck rates made to the S+ pictures, to the 10 test pictures from the same category as the training pictures (reinforced category), and to the 10 test pictures from the previously unseen category (unseen category). The data indicate no evidence of greater generalization to the reinforced category than to the unseen category on the first two test sessions. Thereafter, the rate of pecking to the pictures from the unseen category dropped on Sessions 3 to 5, and the rate of pecking to the reinforced-category pictures and to the S+ pictures remained high.

An ANOVA performed on these data indicated no significant effects of picture condi-

Table 3

Peck rates (pecks per minute) on five test sessions to the S+ pictures, the S− pictures, the novel pictures from the same category aS the S+ pictures, and the novel pictures from the same category as the S− pictures (Experiment 3).

	Test pictures									
	Training S+					Training S−				
Pigeon	1	2	3	4	5	1	2	3	4	5
6095 (car S+)	77.2	68.1	61.1	72.3	76.2	20.8	11.9	12.7	14.7	29.5
4804 (car S+)	47.8	51.6	42.9	31.9	41.3	14.3	26.0	19.0	21.3	16.6
8185 (car S+)	35.0	37.2	30.5	30.8	33.5	19.4	16.5	18.8	11.3	15.1
4 (cat S+)	37.7	44.0	39.7	36.9	41.7	14.8	21.1	13.0	19.3	16.6
693 (cat S+)	97.1	99.0	88.5	93.0	84.7	26.0	25.4	42.2	30.7	25.3
197 (cat S+)	36.5	37.6	37.1	42.2	41.5	14.8	16.1	17.9	20.2	16.7
M	55.2	56.2	50.0	51.2	53.2	18.4	19.5	20.6	19.6	20.0
SEM	9.6	8.9	8.0	9.5	8.0	1.7	2.1	4.1	2.5	2.2

tion, $F(2, 10) = 3.69$, or session, $F(4, 20) = 1.40$. The Picture Condition × Session interaction was significant, $F(8, 40) = 4.50$, indicating the separation of the curves on the final three sessions. One-way ANOVAs were performed across picture conditions on each session. These revealed no effect at Sessions 1 and 2, *F*s < 1.00. Significant effects of picture condition were found during Session 3, $F(2, 10) = 8.73$, and Session 5, $F(2, 10) = 7.85$, but not during Session 4, $F(2, 10) = 3.37$. Newman–Keuls tests comparing means during Sessions 3 and 5 showed that in both cases peck rate to the novel pictures in the unseen category was significantly lower than the peck rate to either the novel pictures in the S+ category or the S+ pictures. Peck rates to the novel pictures in the S+ category and to the S+ pictures did not differ significantly.

These findings fail to support a critical prediction from the theory of stimulus generalization among pictures within a conceptual category. Although pigeons' key pecks were reinforced over 20 sessions in the presence of 20 different pictures within either the cat or car category, they failed to show any difference in response to novel pictures from these categories on the first two sessions of testing. Stimulus generalization theory predicts that pigeons' key pecks reinforced in the presence of cat pictures should occur more often to new cat pictures than to new car pictures upon initial exposure to these test items and that the reverse pattern should be seen for key pecks reinforced in the presence of car pictures. This pattern of differential responding to the reinforced and new categories did eventually appear during Sessions 3 to 5. This delayed appearance of differential responding is exactly what would be expected if pigeons must experience differential reinforcement in order to form categories. During Sessions 1 and 2, pigeons were able to compare nonreinforced exemplars in the new category with reinforced (training pictures) and nonreinforced (test pictures) exemplars from the training category. This opportunity to discriminate pictures from the two categories allowed pigeons to detect the features that distinguish cats from cars and then to begin to respond more to items from the reinforced category than to items from the nonreinforced category.

EXPERIMENT 3

The finding in Experiments 1 and 2 that pigeons did not show differential responding upon initial exposure to novel items from the reinforced and new categories argues against automatic stimulus generalization within categories. One difficulty with this conclusion is that it is based on a null outcome or a failure to find a difference in peck rates between pictures in different categories. It might be argued that the apparatus, procedure, or stimuli used in these experiments do not provide a sensitive test of initial transfer to novel items. This possibility was examined in Experiment 3 by using a more traditional discrimination procedure before testing with novel pictures. Pigeons were trained to discriminate cats from cars by reinforcing peck-

Table 3

(*Extended*)

Test pictures									
Novel pictures from S+					Novel pictures from S−				
1	2	3	4	5	1	2	3	4	5
58.8	50.8	33.6	51.6	74.4	20.8	18.8	14.8	23.6	37.2
57.2	54.4	44.4	26.4	35.6	24.4	34.0	36.0	28.4	38.0
37.6	39.6	30.8	39.2	46.8	26.4	28.8	33.2	26.8	29.2
24.0	27.2	19.6	24.4	30.0	11.6	12.8	7.2	14.8	12.4
80.4	90.0	87.6	74.8	77.2	41.6	48.0	49.2	45.2	41.6
35.6	43.6	26.0	44.8	44.4	15.2	27.6	22.4	26.4	16.0
48.9	50.9	40.3	43.5	51.4	23.3	28.3	27.1	27.5	29.1
7.6	8.0	9.2	6.9	7.4	3.9	4.6	5.7	3.7	4.6

ing to pictures in one category and extinguishing pecking to pictures in the other category. When a discrimination learning criterion was reached, pigeons were tested with novel cat and car pictures interspersed among the training pictures. If pigeons now respond at a higher rate to new pictures in the reinforced category than to new pictures in the nonreinforced category on the first session of testing, it would indicate that the apparatus and stimuli used here are sensitive to immediate transfer of category discrimination.

Method

Pigeons. Six different adult pigeons served as subjects. They had research histories similar to the pigeons in the previous experiments. They were housed and maintained as in the previous experiments.

Apparatus. The touch-screen apparatus was the same as in previous experiments. The 20 cat images and 20 car images used in training were the same as the training images in Experiment 2. The test images were the same 10 cat and 10 car images used in the novel image tests of the previous experiments.

Procedure. All pigeons were pretrained as in the previous experiments. Subsequently, all pigeons received discrimination training. One group of 3 pigeons was trained to discriminate 20 cat images (S+) from 20 car images (S−). The other 3 pigeons received training with the same pictures but with the reinforcement contingencies reversed (i.e., car images were positive and cat images were negative). Training sessions consisted of 120 trials, with each picture presented three times in a random order containing all pictures. Reinforcement followed the first peck after 15 s on S+ images, and S− images were extinguished after 15 s without reinforcement and were followed by the 10-s ITI. During each session, a discrimination ratio (DR) was calculated by dividing the pigeon's pecks to S+ pictures by the sum of its pecks to S+ pictures and its pecks to S− pictures. Pigeons received one training session per day and were trained to a criterion DR of .70 or greater for four consecutive sessions.

Each pigeon began test sessions immediately after reaching the discrimination criterion. Each test session contained 100 trials. Four trials in each successive block of five trials presented images from the training phase (two S+ and two S−) with the same reinforcement outcomes as in training, and each training image was presented twice in a session. One trial in each block of five contained a novel test image. As in the previous experiments, there were 20 novel test images (10 pictures of cats and 10 pictures of cars), and all pigeons saw all test pictures once per session. As in previous experiments, responses to novel test pictures were never reinforced. The position of the novel picture within blocks of five trials was randomized, as was the order of the novel test images across the session. Testing continued for five sessions.

Results and Discussion

Pigeons in the cat S+ group required a mean of 13.33 sessions to reach the discrimination criterion, and pigeons in the car S+

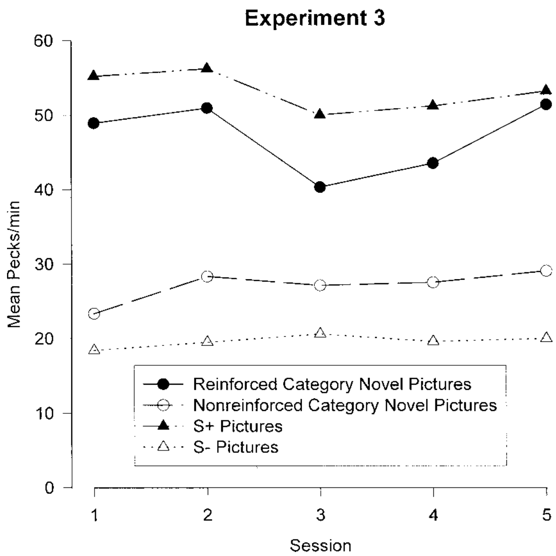

Fig. 3. Peck rates on novel test pictures from the reinforced and nonreinforced training categories plotted over five sessions (Experiment 3).

group required a mean of 18.7 sessions to reach criterion. These means did not differ significantly, $t(4) = 1.12$.

Table 3 shows peck rates of individual pigeons over five test sessions for the training S+ and S− pictures and for the novel pictures from the same category as the S+ pictures and the same category as the S− pictures. Figure 3 shows mean peck rates plotted over five test sessions for each of the picture conditions. Peck rates were considerably higher to S+ pictures and to novel pictures from the S+ category than to S− pictures and to novel pictures from the S− category, both during Session 1 and over all five sessions.

An ANOVA was performed that contained three factors, reinforced versus nonreinforced categories, training versus novel test pictures, and sessions. The only significant effects found were reinforced versus nonreinforced categories, $F(1, 95) = 187.68$, and the interaction of reinforced versus nonreinforced categories with training versus novel test pictures, $F(1, 95) = 12.14$. This interaction arose from the greater difference between S+ and S− pictures than between novel pictures from the S+ and S− categories. Separate t tests showed significant differences between both S+ and S− pictures, $t(5) = 4.29$, and novel pictures from the S+ and S− categories, $t(5) = 4.43$. The same tests performed on the data from Session 1 showed significant differences between mean peck rates to both the S+ and S− pictures, $t(5) = 4.09$, and the novel pictures from the S+ and S− categories, $t(5) = 4.99$.

The finding that pigeons responded more to novel pictures from the reinforced category than to novel pictures from the nonreinforced category during Session 1 indicates that the pictures and procedures used in these experiments were sensitive to the usual effects of transfer to novel items found after discrimination training. When pigeons compared these training pictures in the context of differential reinforcement, categorization of novel pictures appeared during their initial exposure.

GENERAL DISCUSSION

When pigeons' key pecking was reinforced in the presence of either a single picture from a category (Experiment 1) or multiple pictures from a category (Experiment 2), the pigeons showed no greater initial response to novel pictures from that category than to those from a previously unseen category. When the pigeons had an opportunity to compare pictures from two categories, however, differential responding to reinforced and nonreinforced categories appeared. In Experiment 2, pigeons began to peck faster at the training-category test pictures than at the new-category pictures after two sessions in which they could compare the pictures from these different categories. When pigeons were trained with pictures associated with reinforcement and its absence from cat and car categories in Experiment 3, they showed immediate marked differential responding to novel pictures from these categories during Session 1.

These findings appear to challenge the conclusions reached by Herrnstein and his colleagues (Cerella, 1979; Herrnstein & de Villiers, 1980; Herrnstein et al., 1976) that perceptual constraints lead pigeons to detect within-category similarities immediately upon the perception of pictures in the same category. Similarly, they challenge the notion that reinforcing responses to a picture in a given category automatically produces an excitatory tendency to respond to other pictures in the category that are similar to the reinforced stimulus (Astley & Wasserman, 1992). The data suggest that pigeons, and by extension other animals, may begin to detect within-category similarities and differences among categories only when they are forced to discriminate between these categories. Pigeons may indeed store representations of many pictures from a category and may be able to identify these pictures correctly in a recognition test (Vaughan & Greene, 1984), but the current findings suggest that they do not recognize these representations as belonging to a common category unless they discriminate them from items in a different category.

If these conclusions are correct, an interesting question arises: What do animals need to discriminate between category and extracategory images to establish a category concept? Must animals discriminate the members of one category from distinctly different members of another category, or does discrimination from a variety of items in many other categories work just as well? Do pictures in one category have to be discriminated from whole pictures in another category, or would just some features of pictures in another category work as well? Can the degree of category learning, as shown by transfer to novel items, be influenced by the number or type of items shown in another category? If a pigeon has learned that responses to cat pictures are reinforced and those to car pictures are not reinforced, would it also discriminate between novel cat pictures and novel pictures of trees, people, and flowers, or has it learned only that cats are not cars? If the idea presented here, that animals only learn about a category through comparison between categories, is correct, these research questions should be pursued.

REFERENCES

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.

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

Herrnstein, R. J. (1979). Acquisition, generalization, and discrimination reversal of a natural concept. *Journal of Experimental Psychology: Animal Behavior Processes, 5,* 116–129.

Herrnstein, R. J. (1984). Objects, categories, and discriminative stimuli. In H. L. Roitblat, T. G. Bever, & H. S. Terrace (Eds.), *Animal cognition* (pp. 233–261). Hillsdale, NJ: Erlbaum.

Herrnstein, R. J., & de Villiers, P. A. (1980). Fish as a natural category for people and pigeons. In G. H. Bower (Ed.), *The psychology of learning and motivation: Advances in research and theory* (Vol. 14, pp. 60–97). San Diego: Academic Press.

Herrnstein, R. J., & Loveland, D. H. (1964). Complex visual concept in the pigeon. *Science, 146,* 549–551.

Herrnstein, R. J., Loveland, D. H., & Cable, C. (1976). Natural concepts in pigeons. *Journal of Experimental Psychology: Animal Behavior Processes, 2,* 285–302.

Pavlov, I. P. (1927). *Conditioned reflexes.* Oxford, UK: Oxford University Press.

Roberts, W. A. (1998). *Principles of animal cognition.* Boston: McGraw-Hill.

Spence, K. W. (1936). The nature of discrimination learning in animals. *Psychological Review, 43,* 427–449.

Spence, K. W. (1937). The differential response in animals to stimuli varying within a single dimension. *Psychological Review, 44,* 430–444.

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

Wasserman, E. A. (1995). The conceptual abilities of pigeons. *American Scientist, 83,* 246–255.

Wasserman, E. A., & Bhatt, R. S. (1992). Conceptualization of natural and artificial stimuli by pigeons. In W. K. Honig & J. G. Fetterman (Eds.), *Cognitive aspects of stimulus control* (pp. 203–223). Hillsdale, NJ: Erlbaum.

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.

Received August 3, 2001
Final acceptance June 3, 2002