

The Person Concept in Monkeys (*Cebus apella*)

M. R. D'Amato and Paul Van Sant
Rutgers—The State University

In three experiments using person and nonperson slides, 7 monkeys (*Cebus apella*) transferred differential responding significantly to new instances of the person concept on their first exposure to the test slides. The number of exemplars with which the monkeys had experience before the emergence of significant first-trial transfer ranged from 20 to 100. In Experiment 4, 3 of the 7 monkeys were tested with a different collection of person/nonperson slides. Their high rates of transfer indicated that the results of Experiments 1-3 were not limited to the set of slides used in those studies. Despite this strong evidence for representation of the person concept, detailed analysis of the monkeys' errors raised serious questions about the degree to which their transfer behavior was governed by conceptual processes.

The ability of animals to form concepts based on natural categories has been a topic of increasing interest since Herrnstein and Loveland published their pioneering paper in 1964. They reported that pigeons learned to discriminate between slides that contained people from slides that did not and to generalize at a high level to new instances of person and nonperson slides. Most related studies since that time have employed pigeons as subjects and have provided substantial evidence suggesting that this species develops and employs concepts based on such natural categories as bodies of water, trees, fish, and oak leaves, in addition to people (e.g., Cerella, 1979; Herrnstein, 1979; Herrnstein & de Villiers, 1980).

Curiously, in spite of the long interest in concept formation in nonhuman primates (e.g., Andrew & Harlow, 1948), the evidence for natural category concepts in monkeys is less compelling than that amassed for pigeons. In an early study by Lehr (1967), a rhesus and a cebus monkey trained on three exemplars of insects and flowers generalized at a reasonably high level to new instances of the two categories. Apparently, little further work with monkeys was undertaken until recently. Working with stump-tailed monkeys, Schrier, Angarella, and Povar (1984) found only weak evidence for the natural category concepts of people and monkeys. Yoshikubo (1985) provided stronger evidence of the monkey concept in rhesus monkeys, but as discussed below, a methodological problem clouds the interpretation of his results. Finally, Schrier and Brady (1987), using a very large number of exemplars, investigated the person concept in rhesus monkeys and obtained stronger evidence for it in this species.

This research was supported by National Science Foundation Grants 8207146 and 8417383. A preliminary report based on the results from Poe, Spider, and Dagwood appeared earlier (D'Amato & Salmon, 1983).

We wish to thank David Salmon for his help in Experiment 1, and Kathy Whitaker and Beth Borowsky for running the monkeys in Experiment 4. We are grateful to Ben A. Williams for alerting us to the Lehr (1967) article and providing us with an English translation.

Correspondence concerning this article should be addressed to M. R. D'Amato, Psychology Department, Busch Campus, Rutgers University, New Brunswick, New Jersey 08903.

In many of the studies that reported natural category concepts in animals, the subjects were exposed to a large number of different positive and negative exemplars before evidence of concept-mediated transfer was obtained. For example, in the Herrnstein and Loveland (1964) study, the subjects were presented in each session with approximately 80 different exemplars of the target concept, 40 positive and 40 negative. With such a procedure, a subject that displayed substantial transfer as early as the fourth or fifth session would have previously viewed some 240 or 320 exemplars. Given the impressive memorial capacity of pigeons and monkeys (e.g., Ringo & Doty, 1985; Vaughan & Greene, 1984), a considerable amount of positive transfer to new instances could be due to stimulus generalization from specific components of previous exemplars (see Greene, 1983), as opposed to concept-mediated transfer. This interpretive problem, discussed further below, also exists for those studies in which animals were trained on a fixed set of 80 exemplars before being tested for transfer (e.g., Herrnstein, 1979; Schrier et al., 1984; Yoshikubo, 1985).

In an effort to address this issue, we conducted a series of studies with cebus monkeys as subjects, aimed at limiting the number of exemplars used in the training phase while still obtaining high levels of transfer to new instances of the target concept. Although we were reasonably successful in achieving our goal, detailed analysis of the monkeys' responses to novel positive and negative exemplars raised troublesome questions about the degree to which their transfer behavior was mediated by conceptual processes.

Experiment 1

In this study the number of exemplars of the person concept used in the training phase was restricted to 20, 10 positive and 10 negative. Rather than training the monkeys on all 20 exemplars concurrently, which is the usual procedure employed in such studies, we trained them on a sequence of 10 simple discriminations between a single positive and a single negative exemplar. We thought the sequential technique might be more effective than the concurrent procedure in

drawing the monkeys' attention to the target category. A series of transfer tests followed acquisition of the 10 discriminations.

Method

Subjects

Two laboratory-born female *Cebus apella* monkeys, Poe and Spider, approximately 12 years of age, served as subjects. They had a variety of previous experimental experience, but none with the stimulus materials used in this study. They were housed in individual cages, with water constantly available. Food (Purina Monkey Chow 5045) was restricted to a single feeding (0.5–2 hr following an experimental session), adjusted to the maximum amount that supported reliable performance in each monkey. The daily rations were sufficient to maintain the animals at 95%–100% of their free-feeding body weights.

Apparatus

A modified stainless-steel primate cage (70 × 61 × 61 cm) located in a small testing room served as the experimental chamber. An 18 × 18-cm window was cut in one wall of the chamber and covered with a sheet of 3.2-mm clear plastic. A piece of ground optical glass, suspended about 1.5 cm behind the plastic window, functioned as a rear-projection screen. Centered below the window, about 14 cm above the floor of the cage, was a single response lever. Noyes 190-mg banana-flavored pellets could be dispensed to a food cup located on the left side of the window. Ambient illumination in the test room was provided by a reflector-mounted soft-white 25-W bulb, positioned over the center of the test chamber, 1.25 m above its floor. White noise presented from a ceiling speaker served to mask extraneous sounds. Presentation of slides, as well as other trial events and data recording, was accomplished by a Commodore PET 2001 microcomputer.

Stimulus slides were presented by a Kodak 850H autofocus projector, employing a 102-mm f2.8 Ektanar C lens and a 250-W lamp. The projector bulb was maintained at the low power setting, and a Kodak neutral density filter (.70 rating) was placed in front of the projector lens. The maximum height and width of the slides as they appeared on the rear projection screen was 11.5 and 14.5 cm, respectively.

A total of 126 slides was used, most of which were from a private collection of travel slides; almost all were in color. Half contained one or more persons, differing in age, sex, race, and nationality. The extent of exposure of the humans in these person slides varied from head only to full body. The other half of the slides had no features in common other than the absence of human beings; animals were represented in two of these nonperson slides.

Procedure

Training. The monkeys were trained on a sequence of 10 separate go/no-go discriminations. For Poe, the positive stimulus (S+) of each discrimination consisted of a slide in which at least one person appeared; no persons appeared in the negative stimulus (S-) slides. For Spider, this relation was reversed, with nonperson slides serving as S+ stimuli and person slides as S-. During training and testing, both subjects were exposed to the same sequence of slides. A typical training trial proceeded as follows. After a 40-s intertrial period, a slide appeared on the projection screen. If it was an S+ slide, pressing the lever within 5 s of the slide's onset was considered a correct response, resulting in delivery of a food pellet and, after 1 s, termi-

nation of the slide and entry into the intertrial interval (ITI). Failure to press the lever within 5 s was an incorrect response and resulted in termination of the slide and immediate entry into the ITI. On an S- slide presentation, withholding responding to the lever for 5 s was a correct response, resulting in termination of the slide and entry into the ITI. Responding within 5 s was an error, postponing termination of the slide until 15 s had elapsed from the last lever press. Note that the same contingency and consequences were associated with an incorrect S+ trial and a correct S- trial. In both cases a nonresponse period of 5 s was followed by termination of the slide and entry into the ITI.

A training session consisted of 40 trials, with equal numbers of S+ and S- trials randomly intermixed, subject to the restriction that no more than three of the same trial type could occur in succession. Usually, only one session was given in a single day. Training on each discrimination continued until the subject met the joint criteria of 75% or more correct responses in a session and significantly (.05 level) shorter response latencies on S+ than on S- trials.

Transfer tests. In the first transfer test, the subjects were exposed to 40 slides during a single session, the first 20 of which were the 10 S+ and the 10 S- slides of training (old slides). The second 20 consisted of slides that the monkeys had never seen before (new slides), 10 person and 10 nonperson slides. Within the blocks of old and new slides, the stimuli were arranged in a random order, subject to the restriction that no more than three person or nonperson slides could occur in succession. The usual reinforcement contingencies were in effect during all transfer tests.

There were four additional transfer tests, each consisting of 20 old and 20 new slides, half person and half nonperson slides. The old slides were drawn from previous test or training slides; the new slides were, of course, novel slides. Unlike Test 1, in Tests 2–5 old and new slides were randomly intermixed, with the restrictions that the first two slides were always old and no more than three person or nonperson slides could occur in succession. Before proceeding from one transfer test to the next, the monkeys were trained on the current set of test slides until their performance on the new slides reached a reasonably high level, which occasionally required discrimination training on specific pairs of new slides; the same procedure was followed in Experiments 2 and 3.

When Test 2 was completed, in an effort to increase her S- response latencies, Poe was retrained on familiar slides with correct S- trials now food reinforced; this modified S- contingency remained in effect during Tests 3–5.

After completion of Test 3, the monkeys were trained on three new discriminations based on matched pairs of slides, which were created by photographing the same scene twice, once with one or more persons present and once with people absent. In some cases an object (e.g., a box, plant, or chair) replaced the person in the matching positive slide. The point of this training was to prepare the subjects for the 20 new slides of Test 4, which were arranged into 10 such matched pairs. Test 5 included 5 matched pairs among the 20 new slides.

Results and Discussion

Training

The total number of sessions required to reach the acquisition criteria on all 10 discrimination problems was 15 for Poe and 34 for Spider. Rate of acquisition increased dramatically over the 10 discriminations. On the last five problems, Poe satisfied the learning criteria in the minimum of a single session; Spider required two sessions on three problems and

only one on the other two. Although the accuracy criterion was set at 75% correct, Poe and Spider averaged 92.5% and 90.0% correct responses, respectively, on the criterial sessions of the last five discriminations.

Transfer Tests

Figure 1 presents the percentages of correct responses on old and new slides for the first two sessions of Tests 1-5. The criterion we employed to indicate significant transfer to the new slides was at least 75% correct responses on these slides during a single session (solid line in Figure 1), which by the binomial distribution has a two-tailed probability level of .042. (The two-tailed probability of achieving 80% correct responses by chance is .012.) Inasmuch as the normal reinforcement contingencies were in effect during the transfer tests, the interpretation of significant transfer on Session 2 is complicated by the learning opportunity provided by Session 1.

On Tests 1-2, significant transfer to the new slides was observed only in Spider on the second session of Test 2. However, both animals showed significant transfer to the new

slides on the first session of Test 3, with Poe performing at the same level on old and new slides.

The new slides of Test 4 were matched pairs, which may account for the low performance level observed during the first test session. Poe responded at a very high level of accuracy to the new slides on the second session, while Spider's performance improved only slightly. Despite the fact that 10 of the 20 new slides of Test 5 were composed of matched pairs, the 2 subjects showed significant transfer both on the first and second sessions of this test, Spider performing at the same level on old and new slides in both sessions. It might be noted that pigeons tested with similar matched pairs of slides also showed significant transfer (Siegel & Honig, 1970); however, the birds had previously been trained on a set of 280 person and nonperson slides.

Performance on the new slides was also evaluated by Mann Whitney *U* tests, which were based on the response latencies of the 10 S+ and the 10 S- trials. As might be expected from the fact that correct S- responses required that the subjects withhold responding for 5 s, response latency was a somewhat more sensitive indicator of transfer than was response accuracy. There were three cases in which significant transfer was obtained with the latency measure but not with response

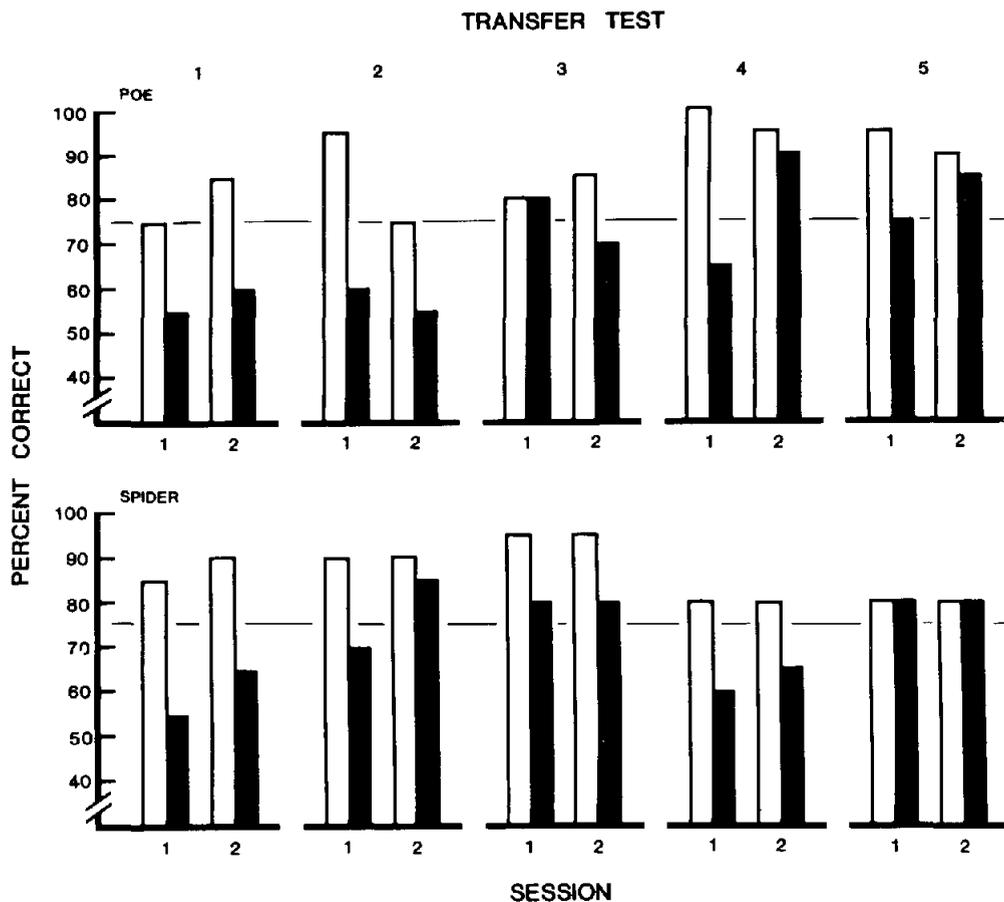


Figure 1. Percentage of correct responses on old and new slides (open and filled bars, respectively) during the first two sessions of the five transfer tests of Experiment 1. (The solid line indicates the .05 level of significance.)

accuracy, and only one in the reverse direction. We will adopt the more conservative percentage of correct responses as our index of transfer and not consider response latency any further.

To summarize the results of the transfer tests, significant transfer on the first test session was obtained in both subjects after previous experience with only 30 person and 30 nonperson exemplars—a rather notable outcome, given the numbers of training exemplars used in many earlier studies. Significant transfer on the first test session reappeared in Test 5, even though half of the new slides were arranged in matched pairs. Although person slides served as positive stimuli for Poe and as negative stimuli for Spider, there was little difference in the degree of transfer displayed by the 2 monkeys.

Experiment 2

The training procedure employed in Experiment 1 and those used in most previous relevant studies may not be very efficient in promoting expression of natural category concepts in animals. The critical features of positive exemplars that differentiate them from negative exemplars must be inferred by the animal from a host of other variable features. Also, in pigeons at least (Greene, 1983), irrelevant background cues tend to gain discriminative control. As has been suggested before (Siegel & Honig, 1970), categorical responding might be improved if matched pairs of positive and negative exemplars, differing primarily in the presence or absence of the target stimulus, were used in training. In the present study we explored this possibility.

Method

Subjects and Apparatus

Two adult female *Cebus apella* monkeys (Dagwood and Olive) and 1 male (Roscoe) served as subjects; all were wild-born and approximately 20 years of age. They had a great deal of previous experimental experience but none with the stimulus materials of the present study. Housing and maintenance were as in Experiment 1, as were the apparatus and slides.

Procedure

Training. The major change from Experiment 1 was that 5 of the 10 discriminations were based on pairs of slides that were very similar except that the positive exemplar contained one or more persons; in all cases, person slides served as the S+ stimuli. Three of the five matched-pair discriminations employed the training slides used in Experiment 1 prior to Test 4; the other two were based on matched pairs of slides drawn from Test 5. For Dagwood and Olive, the last five discriminations were based on matched pairs; for Roscoe, they were discriminations Nos. 4, 5, and 8–10. The training parameters were generally the same as in Experiment 1. The acquisition criteria were strengthened, however; 85% correct responses and 10 correct responses in succession within the same session were required.

Transfer tests. In an effort to improve transfer performance on Test 1, 12–14 pretest sessions were given in which the 20 training slides were presented in a quasi-random order twice during each session. The point of these sessions was to familiarize the subjects

with the multiple-slide procedure used in the transfer tests. Because Roscoe performed poorly on the initial pretest sessions, the penalty for an incorrect S– response was changed to a 60-s time-out, signaled by extinguishing both the slide and the overhead light; the ITI was reduced to 20 s. These modifications greatly improved his performance.

Transfer Tests 1 through 4 employed the same new slides as in Experiment 1, the sole exception being that Roscoe in Test 2 was given one person and one nonperson slide that had served as training slides for the other subjects. Apart from the substitutions for the two matched pairs of slides that were used in training, the new slides of Test 5 were also the same as in Experiment 1.

Olive, who started the study before Roscoe, was advanced to Test 1 even though her performance on the pretest sessions was poor. She was removed from the experiment after the initial test. Three months later she was returned to the pretest phase; the same parameters were used that had proved successful with Roscoe.

Results and Discussion

Training

The total numbers of training sessions required to meet the acquisition criteria on the 10 discriminations were 33, 42, and 50 for Dagwood, Roscoe, and Olive, respectively. The monkeys learned the matched-pairs discriminations very rapidly, averaging only 1.67 sessions to satisfy the criteria on the last three discriminations, which for all subjects were matched pairs.

Transfer Tests

Figure 2 presents the transfer results for the first two sessions of each test. Olive's Test 1 results are based on data obtained before her 3-month break. Despite the pretest training with multiple slides, none of the monkeys transferred significantly to the new slides of Test 1 on the first session.

Training with matched pairs, on the other hand, seemed to have its intended effect. Roscoe showed a high degree of transfer on Tests 2–5, managing at least 80% correct responses on the first session of each test, and Olive transferred significantly on the first session of Tests 3–5. Both subjects also showed significant transfer on the first session of Test 4, in which the new slides were matched pairs, whereas in Experiment 1 neither Poe nor Spider did so. Because significant first-session transfer is often taken as evidence for concept-mediated transfer, we may say that Roscoe and Olive produced more evidence of such transfer than did the 2 subjects of Experiment 1. Still, the fact that the transfer performance of Dagwood was inferior to that of Poe and Spider indicates that although training with matched pairs appears to promote evidence of concept-mediated transfer, it does not by itself assure such a result.

Experiment 3

Expression of conceptual behavior in a first-trial transfer paradigm tends to be fragile, and it may be suppressed or facilitated by a variety of procedural variables. For Roscoe and Olive, extending the duration of S– was not as effective

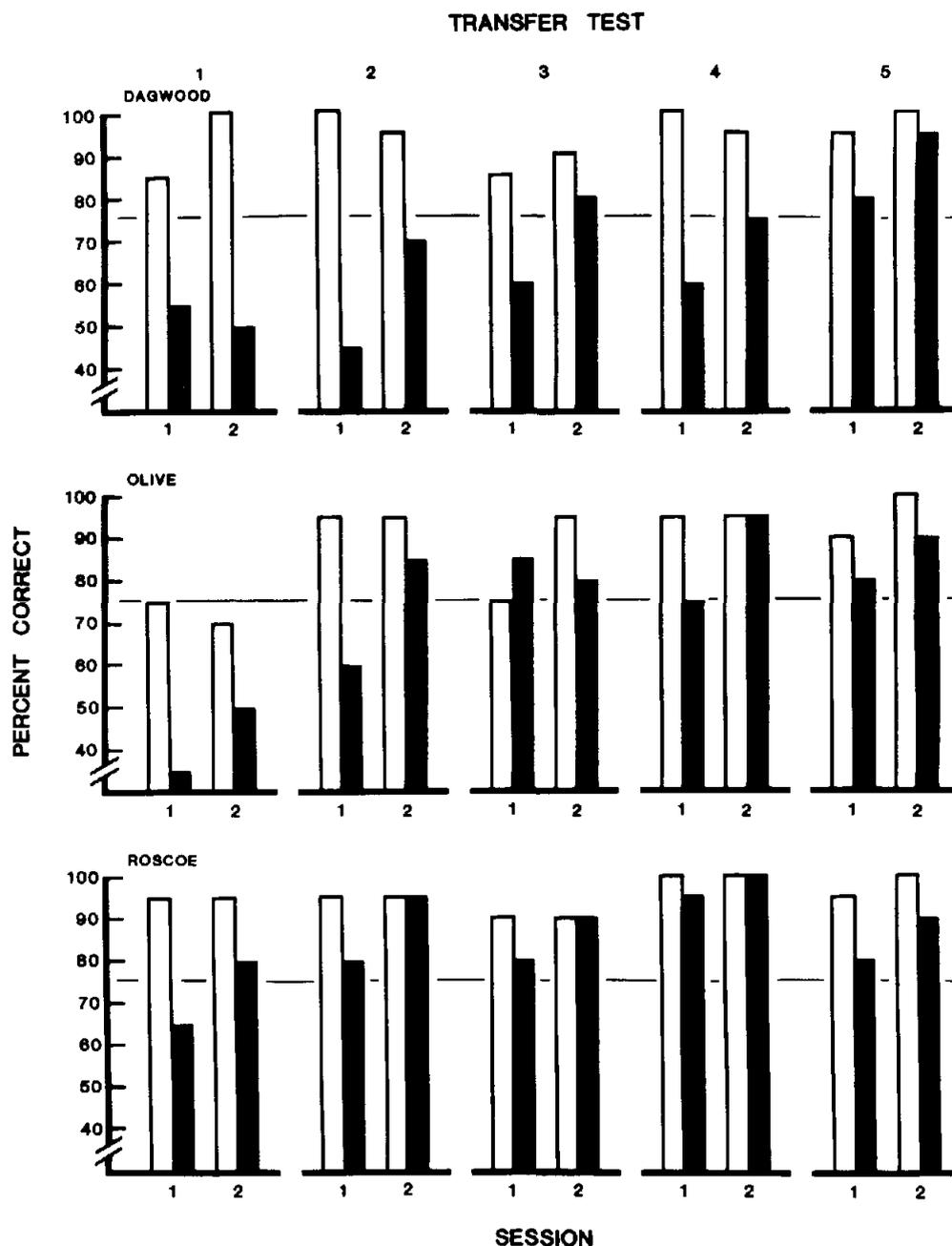


Figure 2. Percentage of correct responses on old and new slides (open and filled bars, respectively) during the first two sessions of the five transfer tests of Experiment 2. (The solid line indicates the .05 level of significance.)

a punishment for incorrect responses on S- trials as presenting a 60-s time-out, possibly because of the long history our monkeys have had with the latter contingency. In an effort to reduce further the number of previously experienced exemplars required for significant transfer, in the present experiment 2 cebus monkeys were trained from the outset with the time-out contingency on incorrect S- trials.

A second aim of the study was to assess whether skin color is an important defining feature of the person category as conceived by our monkeys within the present context. The

monkeys were trained and tested with person slides that contained only individuals having light skin color. On the fifth transfer test, S+ slides were introduced that contained one or more Black persons.

Method

Subjects and Apparatus

One adult female *Cebus apella* monkey (Coco) and one male (Pete) served as subjects; both were wild-born and approximately 20 years

of age. Coco had received substantially more previous experimental training than had Pete, but neither animal had any experience with the stimulus materials of the present study. Housing and maintenance were as in Experiment 1. The apparatus and slides were the same as employed in the previous experiments.

Procedure

Training. The sequence of 10 discrimination problems was very similar to that employed with Dagwood and Olive in Experiment 2: The same discriminative stimuli were used in 7 problems; four of the five matched-pair discriminations employed the same slides and occurred in the same order, the fifth being replaced with a different matched pair to avoid an S+ slide that contained a Black person. Except for the 60-s time-out on incorrect S- trials and reduction of the ITI to 20 s, the training parameters were the same as in Experiment 2.

Transfer tests. The new slides employed in Tests 1-4 were the same as used with Dagwood and Olive except for the substitution of 7 S+ slides that contained Blacks or dark-skinned persons with slides containing only Whites. Three of the 10 new S+ slides of Test 5 contained one or two Blacks, and another three contained one or two Blacks along with one or two Whites. Of the 20 new slides, 12 were the same as used in Test 5 of Experiment 2, while the remaining 8 were slides that had occurred either in training or testing in Experiment 2.

As in the previous experiment, pretest sessions (6 for Coco, 2 for Pete) based on the 20 training slides were given to familiarize the subjects with multiple-slide sessions. In the first session of Test 1, the 20 old slides preceded the 20 new slides; the testing procedures were otherwise the same as in Experiment 2.

Results and Discussion

Training

Each of the 2 subjects required only a total of 24 sessions to reach criterion on the 10 training discriminations. This value is 27% below the smallest total number of sessions (33) observed in Experiment 2 and 42% below the mean total number of sessions (41.7). It appears, therefore, that the training parameters were generally more efficient than those used in the previous study. However, there was little difference in the rate at which the last three, matched-pair, discriminations were learned, 1.83 vs. 1.67 sessions in Experiment 2.

Transfer Tests

Figure 3 presents the results of the five transfer tests. Pete responded at a very high level of accuracy on the first session of Test 1, the only subject to provide evidence of concept-mediated transfer on the initial transfer test. In Sessions 1 and 2 of Test 2, this subject managed 95% correct responses on both the old and the new slides, another first. His inexplicably poor performance in Test 3 is the result of a motivational deficit, possibly arising from a minor ailment. Pete appeared very lethargic in the testing cage during both sessions, and he responded only once in each session to the new slides. It was more than 3 weeks, during which time he received 20 sessions of retraining on previous test and training slides, before his

behavior recovered to the point that he could be advanced to Test 4.

Pete again produced significant transfer on the first session of Test 4, which, as previously, was based on matched-pairs of slides. Although he failed to transfer significantly on the first session of Test 5, five of the six errors on new slides committed by him occurred during a period of loud commotion caused by the escape of an animal from its cage in the colony room, which was located near the testing room. As if in compensation, Pete scored 100% correct responses on the new slides in Session 2. The first evidence for concept-mediated transfer appeared in Coco on Test 4. Her performance on the new slides was very high in Test 5, 90% correct on Session 1 and 95% on Session 2.

Black persons were represented for the first time in six of the 10 new S+ slides of Test 5, Blacks alone appearing in three of the slides. Neither monkey made a single error on any of these slides during Sessions 1 and 2. On the other hand, four errors were committed on the other four new S+ slides. Within the present context, therefore, skin color does not appear to be a critical defining feature of the person category.

With respect to the relative effectiveness of the 60-s time-out contingency associated with incorrect S- trials, our impression is that it facilitated training and probably had a positive effect on transfer as well.

Experiment 4

All 7 monkeys in Experiments 1-3 provided evidence of concept-mediated transfer during the testing series, and 5 subjects did so after exposure to only 60 or fewer exemplars of the person concept, which is rather impressive. However, the same set of slides was used in all three studies. It could be argued that our results might be largely due to insufficient variation among the person slides, that is, essentially to stimulus generalization from specific previous exemplars. Although this interpretation seems strained, particularly in view of the matched-pair slides used in Test 4 of each experiment, a new set of 60 person and 60 nonperson slides was obtained from another investigator, who had used the slides in similar research with rhesus monkeys (Schrier & Brady, 1987). Three subjects from the previous studies were tested with these new slides. If the transfer results that had been obtained with these monkeys were largely due to a high level of similarity among the person slides, one would expect that significant transfer would not be observed with the new slide set, at least not on the first one or two transfer tests.

Method

Subjects and Apparatus

Coco, Dagwood, and Pete, who participated either in Experiment 2 or 3, served as subjects. The housing, maintenance, and apparatus were the same as in the previous experiments.

Most of the 120 slides of the new slide set were based on photographs of scenes from magazines. Because the target scenes usually did not completely fill the slides, which often contained text material

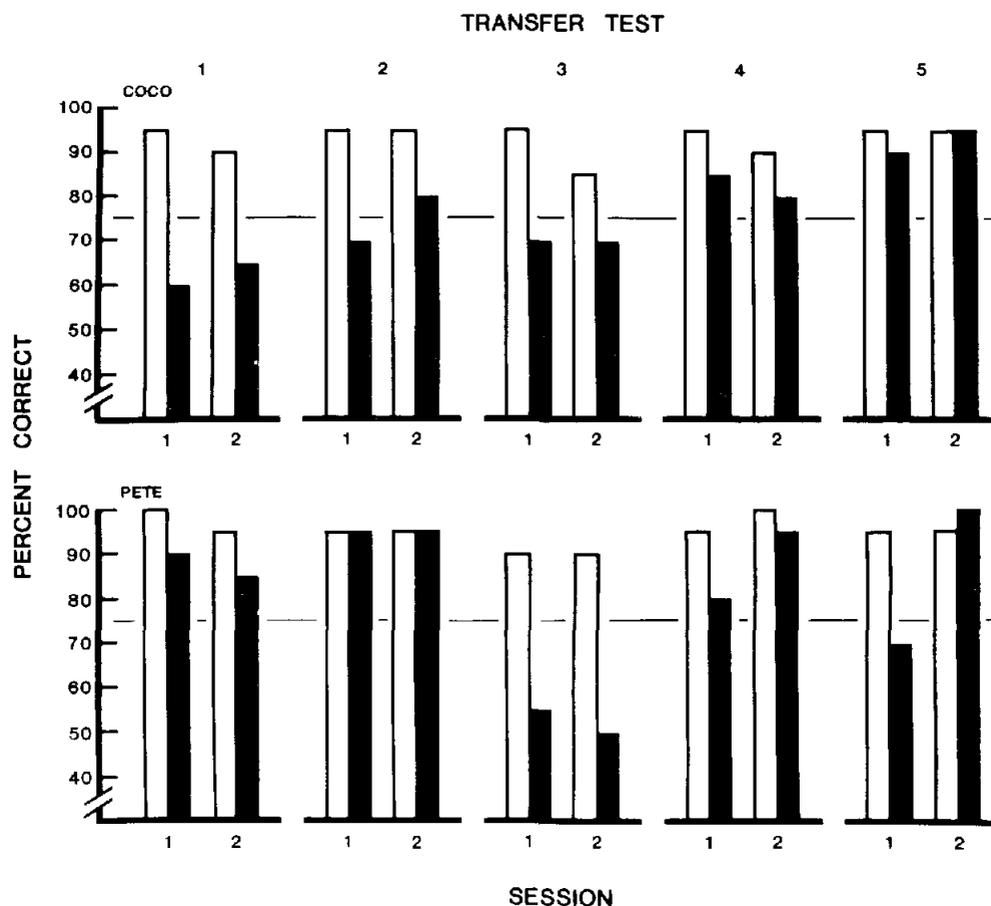


Figure 3. Percentage of correct responses on old and new slides (open and filled bars, respectively) during the first two sessions of the five transfer tests of Experiment 3. (The solid line indicates the .05 level of significance.)

as well, opaque tape was used to reduce the size of the projection screen to 9.3 cm in height and 7.5 cm in width.

Procedure

Retraining. Because the monkeys had been away from the task for from 9 to 14 months, they were retrained on the person/nonperson discrimination with 40 of the slides employed in the previous studies. After reaching an accuracy level of at least 90% correct, 10 person and 10 nonperson slides were replaced with other slides from the earlier slide set, and training was continued until the same criterion was met; reduction of the projection screen took place during this stage of retraining. Finally, another such substitution was made, and the animals trained until the 90% criterion was satisfied. Where a subject had continuing difficulty with a particular slide, it was given one or more sessions with explicit discrimination training on this and a companion slide. The total numbers of retraining sessions ranged from 12 (Dagwood) to 24 (Coco). The trial procedures and contingencies during retraining and during transfer testing were identical to those of Experiment 3.

Transfer tests. There were four transfer tests, conducted in same manner as previous tests. The old slides of Test 1 were drawn from the retraining slides; in Tests 2-4 they always were the new slides from the immediately preceding test. To facilitate analysis of performance on individual slides, all tests were continued for five sessions.

For the most part, the new slides of Test 1-3 were selected in accordance with the numbered order that was assigned by the supplier. Those of Test 4 were chosen from the remaining usable slides so as to include a number of what we considered to be relatively poor exemplars of person slides (Figure 4, top).

Results and Discussion

The results from the first two sessions of each transfer test are shown in Figure 5. With the sole exception of Dagwood on Test 1, evidence of concept-mediated transfer was obtained in all subjects on all tests. Pete's significant transfer on the first session of Test 2 is particularly noteworthy. This subject, who seemed to attract misfortune, experienced essentially random reinforcement during most of the session. Owing to a momentary loss of power to the slide projector on Trial 8, slide presentation fell out of synchrony with the computer program, so that beyond Trial 8 correct S+ responses were sometimes food-reinforced and sometimes punished with the time-out. Although 19 of the 20 new slides appeared after Trial 8, Pete managed to achieve a performance level of 75% correct on these slides. To restore his confidence in the consistency of the contingencies, Pete was given two sessions



Figure 4. Top: One of the poor exemplars of person slides used in Test 4 of Experiment 4. Center: A troublesome nonperson slide of Test 3 in Experiments 1–3; a patch of red was produced by the piece of watermelon on the table. Bottom: A persistently misclassified nonperson slide of Test 1, Experiment 4; two bright red spots appeared on the animal's body. (The three slides were in color.)

with the slides of Test 1 before being advanced to Session 2 of Test 2.

The significant transfer observed in all subjects on the first session of Test 4 is impressive in view of the fact that poor

exemplars of the person concept were included in the new person slides. The results of this study clearly indicate that the transfer obtained in Experiments 1–3 was not significantly dependent upon the particular slide set used in those studies.

General Discussion

Evidence for concept-mediated transfer, defined as significant transfer to the new slides on the first session of a test series (which required an accuracy level of at least 75% correct), was displayed by all 7 subjects in Experiments 1–3. It first appeared in 1 subject after exposure to only 20 exemplars of the person concept, 10 positive and 10 negative instances. One subject required experience with 40 exemplars before the first evidence of concept-mediated transfer appeared; 60 exemplars were needed in three cases, and the remaining 2 subjects succeeded after 80 or 100 exemplars. A different set of slides was employed in Experiment 4, and the fact that significant transfer was observed on the first session in 11 of the 12 transfer tests administered to the 3 subjects indicates that the positive results obtained earlier were not restricted to the particular set of slides used in Experiments 1–3. Given the transfer criteria, explicit or implicit, employed in previous relevant studies, the present results could be regarded as rather convincing evidence for representation of the person concept in cebus monkeys.

Compared with similar studies employing monkeys, our results are considerably stronger than those reported by Schrier et al. (1984, Experiment 1) for stumptailed monkeys. After being trained on a set of 80 exemplars of the person concept, their monkeys averaged only 53.3, 61.7, and 65.0% correct responses on the first three exposures to a new set of 80 slides. More impressive results were obtained by Schrier and Brady (1987) in a subsequent study that employed a very large stimulus set. In 80-trial sessions, rhesus monkeys were presented on each trial with a person and a nonperson slide and required to respond to the former. A critical feature of this study was that different slides were used on every trial of every session. The 7 best performers of their 9 subjects averaged about 63% correct responses on the very first session, which was significantly better than chance; performance over the first 30 trials (60 exemplars) was not above chance.

By way of comparison, the 7 monkeys of Experiments 1–3 averaged 59.3% correct responses during their first exposure to the 20 new slides of Test 1, $t(6) = 1.49$, $p > .05$. However, they averaged 68.6% correct responses to new slides on the first session of Test 2, $t(6) = 3.07$, $p < .05$. Thus, as a group, our monkeys had experienced 40 exemplars before they responded above chance on their first exposure to new slides, which is very similar to the results obtained by Schrier and Brady in the first session of their study.

It appears, however, that our subjects showed faster improvement over the subsequent three transfer tests. In Tests 3–5, which involved a total of only 60 new slides, the mean percentages of correct responses to new slides on the first test sessions were 72.9, 74.3, and 79.3, respectively ($t_s = 5.26$, 4.81, and 12.75; $p < .001$ in all cases). Over the 80 trials (160 slides) of their second session, Schrier and Brady's monkeys

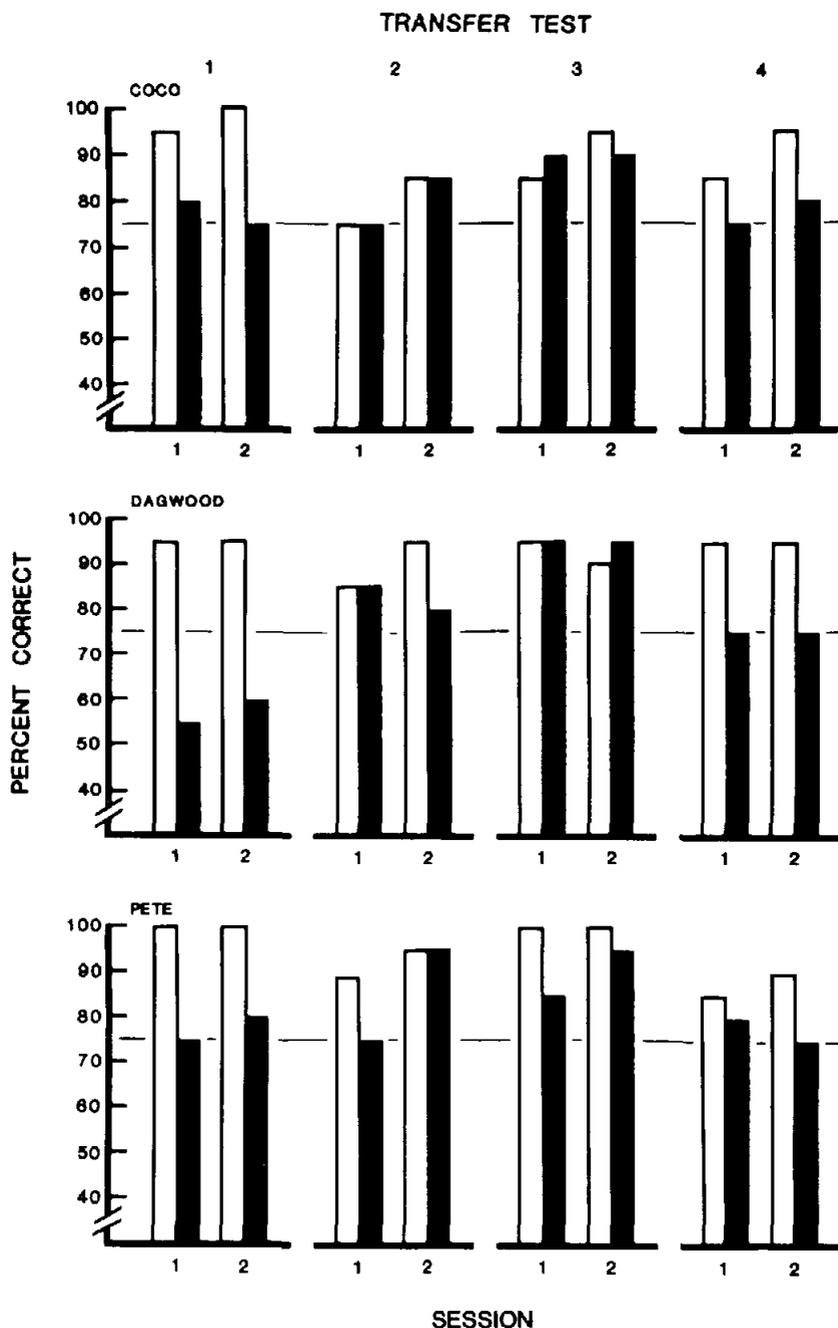


Figure 5. Percentage of correct responses on old and new slides (open and filled bars, respectively) during the first two sessions of the four transfer tests of Experiment 4. (The solid line indicates the .05 level of significance.)

again averaged about 63% correct responses, reaching about 72% correct on the third session.

Still, in terms of numbers of training sessions, Schrier and Brady's procedure seems very efficient. Only a couple of sessions beyond pretraining were required to produce significant group results. And by the end of only 14 sessions, performance reached a level of about 90% correct responses. However, the monkeys had been exposed to 1,120 positive and 1,120 negative instances of the person concept over the 14 sessions. As noted above and in a related context (D'A-

mato, Salmon, & Colombo, 1985), the opportunity for stimulus generalization from specific features of past exemplars increases with increasing numbers of exemplars of the target concept.

It is not possible to compare our results with those of Yoshikubo (1985), who investigated acquisition of the monkey concept in rhesus monkeys. Although his subjects transferred to new exemplars after exposure to 80 or 100 exemplars, all of his transfer results are based on pooled data from the first three exposures to the test stimuli. As was mentioned above

(and is evident in Figures 1–3 and 5), such a procedure provides a learning opportunity that could contaminate the transfer results.

We have noted that in studies of natural category concepts, successful transfer to new exemplars of the target concept could be due to stimulus generalization from specific components or features of previous instances. This assertion raises the very difficult issue of distinguishing between stimulus generalization and concept-mediated transfer, a problem that cannot be addressed in sufficient detail here. Nevertheless, we would like to point out that insofar as transfer to positive instances is concerned, it is very difficult to disentangle the contributions of the two processes, because both are rooted in physical similarity. Transfer to new positive exemplars might arise from the animal's abstracting a variety of relevant features from previous positive instances and assembling them into an abstract representation, such as a prototype, which would qualify as concept-mediated transfer. Or it might be due to "mindless" generalization from a specific relevant or even irrelevant feature of previously encountered positive exemplars. Such considerations suggest that analysis of where transfer fails to occur, that is, of errors, might prove more informative.

Analysis of Misclassifications

Schrier and Brady (1987) classified their person slides beforehand into five subcategories, ranging from poor instances of the person concept (humans comprising less than 25% of the slide) to good instances (at least 50% of the slide devoted to humans). The mean percentage of correct responses over the last nine sessions was monotonically related to the "goodness" of the slides as defined by Schrier and Brady. Only 66% correct responses were obtained with slides of the poor category, whereas good-category slides produced 93% correct responses. If their monkeys' performance was based on a well-differentiated person concept, it is difficult to understand why it should be so tightly coupled to the amount of area of the slide occupied by humans.

The high error rate generated by the poor-category slides is especially perplexing. One might argue that because humans comprised such a small part of the scene in these slides, the monkeys' might simply have terminated their search before detecting them. Even so, why should they then respond to the comparison slides, which were completely devoid of humans? Moreover, Schrier and Brady suggested that for most of these slides, human subjects would have had little difficulty making such detections, a point for which some documentation exists (Schrier et al., 1984). In short, in spite of the high level of transfer displayed by Schrier and Brady's subjects on the later training sessions, the very large number of exemplars to which they had been exposed and the nature of their misclassifications raise some question as to the conceptual basis of the monkeys' transfer behavior.

Analysis of misidentifications evoked similar reservations about our own results. We, too, generally found better performance on slides that would qualify as intermediate to good by Schrier and Brady's criteria. Analysis of performance on selected individual slides proved especially informative. Be-

cause Spider was trained to respond to nonperson slides, she was excluded from the following analyses.

Focusing our attention on Experiments 1–3, during the first two sessions of Transfer Test 2, the 6 monkeys committed a total of only 12 errors on the 10 new person slides of that test. Ten of the 12 errors were made to the same slide, which was a three-quarters frontal view of the head and shoulders of an adult male that occupied most of the scene. To any human observer, this "portrait" slide would constitute an incontestable positive exemplar of the person concept. It clearly was not for the monkeys, perhaps because the slide was rather different from all previous person slides seen by them: It was black/white and, more important, the only person slide that presented a human in such a close-up view. [Interestingly, pigeons classifying fish and nonfish slides also had considerable difficulty with a close-up view of the head region of a fish (Herrnstein & de Villiers, 1980.)]

Recall that the consequences of an incorrect S+ trial were identical to those of a correct S– trial. In spite of the fact that no differential information was provided on incorrect S+ and correct S– trials, some subjects corrected themselves on the troublesome portrait slide, as if they somehow "realized" their previous errors. For example, Dagwood and Olive (Experiment 2) received four sessions in Test 2, and although Dagwood failed to respond to the portrait slide on the first two sessions, as did Olive on the first, both subjects responded correctly on all subsequent sessions. These response reversals in the absence of feedback from previous errors might suggest that in the case of the portrait slide, expression of the person concept was masked on the initial sessions of Test 2, owing to the fact that this slide was so different from previous positive exemplars.

Another interpretation is possible, however. Coco (Experiment 3), who also received four sessions in Test 2, failed completely to respond to the portrait slide. As it turned out, one of the new person slides that somewhat resembled the portrait slide was presented to Dagwood and Olive in Test 2 but, because of the aims of Experiment 3, not to Coco. This slide, also black/white, contained a close-up view of a woman, along with several more familiar instances of the person category. Dagwood and Olive responded correctly to this slide on all but one occasion (Olive in Session 3), and it seems plausible that their change in classification of the portrait slide was largely due to generalization from the former slide to the latter. Interestingly, Pete, who was the other subject in Experiment 3 and therefore, like Coco, was not exposed to the second close-up slide, committed only two errors on the new slides during the two sessions of Test 2 that he received, and both were on the portrait slide.

A number of misclassifications of nonperson slides seemed perplexing until detailed analysis of individual slides appeared to identify a common thread. The nonperson slide shown in Figure 4 (center), which appeared in Test 3 of Experiments 1–3, was surprisingly troublesome. Omitting Pete, who rarely responded to the new test slides during Test 3, a total of 19 sessions was given to the 5 subjects in this test series. In spite of the punishment associated with incorrect S– trials, not a single correct classification of the slide was made by any subject.

We suspect that these persistent errors were elicited to a large degree by a patch of reddish coloration produced by the piece of watermelon in the center of the table. A variety of patches of red, produced mainly by items of apparel, occurred in about a third of the previous person slides, and, in fact, there were two such instances among the other new person slides of Test 3.

A similar persistent misclassification occurred in Experiment 4. Recall that five sessions were administered to each of the 3 subjects on all of the four transfer tests. In Test 1, Coco and Dagwood responded to the new nonperson slide shown in Figure 4 (bottom) on all five presentations; Pete misclassified it on the first two sessions, correcting himself thereafter. Although there is no apparent similarity between a human and the marine creature shown in the slide, the animal's coloration included two bright red patches, one in the head region and one on the body.

To examine further the control exerted on the monkeys' classification behavior by patches of red coloration in nonperson slides, the 40 novel nonperson slides of Tests 1-4 used in Experiment 4 were classified in accordance with whether or not such a patch was present. Of the 13 slides in which this stimulus component was present, the reddish patch was a feature of an animal in three cases, it was a feature of a flower in another three instances, and in the remaining 7 slides it was associated with inanimate objects. Twenty-six of the 40 nonperson slides did not include anything resembling a red patch; finally, one slide that was difficult to classify was eliminated from the analysis.

Summing across the five sessions that each slide was presented, the mean numbers of errors committed on the two types of nonperson slides were calculated for each subject. In the case of nonperson slides lacking a red patch, these values were 0.27, 0.69, and 0.23, for Coco, Dagwood, and Pete, respectively ($M = .40$). The corresponding values for the slides in which the red patch was present were 1.54, 1.38, and 1.38 ($M = 1.43$). Thus, about three and one-half times more errors were made on the nonperson slides that contained a reddish patch; a correlated t test showed the difference between the means to be statistically significant, $t(2) = 5.86, p < .05$.

Examination of the errors produced by individual slides suggested that the effectiveness of the red patch in eliciting responses to nonperson slides depended upon the nature of the responsible agent. Consequently, the errors associated with each category of such slides—animals, flowers, or inanimate—were separately calculated for each subject. The group averages for the three categories were 3.45, 1.33, and .62, respectively; a randomized-blocks analysis of variance indicated that the means differed reliably, $F(2, 4) = 17.72, p \approx .01$. Thus a nonperson slide containing a reddish patch was considerably more likely to elicit errors if the color was a feature of an animal or a flower; still, the error rate for the inanimate category was more than 50% higher than that of nonperson slides which did not contain a red patch.

Feature Versus Concept-Mediated Transfer

The control exerted on the monkeys' discriminative behavior by the red patch component, particularly when it consti-

tuted a feature of an animal or a flower, does not necessarily imply that the significant transfer frequently observed on the initial test sessions of Experiments 1-4 totally lacked a conceptual basis. For one thing, as suggested by others (e.g., Greene, 1983), control by specific local features could coexist with control by conceptual processes. Second, although the topic has been debated in some detail (e.g., Lea, 1984), it is by no means clear how categorization behavior that is mediated by specific features, or even by specific exemplars, is to be distinguished in animals from categorization based on more conceptual mechanisms. Indeed, Medin and Smith's (1984) recent review leaves one with the impression that there is not much agreement as to the mechanisms by which humans form natural concepts or even as to the proper definition of such concepts (see also Shank, Collins, and Hunter, 1986).

Despite the lack of such a consensus and notwithstanding the absence of agreement as to the actual or appropriate role of feature matching in concept-based categorization, there are outcomes that could be taken to argue against a conceptual basis for object categorization behavior, however the latter is defined. For example, consistent failure to transfer to new exemplars would automatically disqualify the process by which the original set of exemplars was learned from having a conceptual basis.

Persistent reliance on irrelevant features of the target concept is also cause for serious concern. Although it is true that natural category concepts ordinarily do not have a set of defining (necessary and sufficient) characteristics (Medin & Smith, 1984), for concept formation to mean something more than discrimination learning there must be constraints on the nature of the features utilized by the subject to categorize positive and negative instances of the target concept. One such constraint surely is that irrelevant features ought not to exert strong control over the subject's categorization behavior. Defining an "irrelevant" feature precisely is not an easy matter, but a useful starting point is *any feature that is not a necessary property or characteristic of any member of the target category*.

One might be willing to allow some control for irrelevant features that occur fortuitously in association with the target concept. In the present studies, for example, because items of red apparel frequently occurred in conjunction with humans, one might expect that new positive exemplars in which humans wore red clothing would be responded to more rapidly than positive exemplars in which this feature was absent. But when an irrelevant feature occurs in conjunction with a clear negative instance and the subjects insist on classifying it as a positive instance, one cannot help but doubt the conceptual basis of their successful classifications.

As a last example, one of the new nonperson slides of the last test of Experiment 4 presented a full view of a jackal carrying a dead flamingo in its mouth; jackal and prey also appeared in a reflection from the shallow water through which the predator was marching. In spite of the considerable experience that the 3 monkeys in the experiment had in classifying person and nonperson slides and the fact that incorrect responses on S- trials were punished with a 1-min time-out, 2 of the monkeys classified the slide as a positive instance five

times in succession, and the 3rd subject classified it correctly for the first time on the fourth trial, only to respond to it again on the fifth presentation! Apparently, the conjoint features of red patch/animal comprised an irresistible instance of the "category" that controlled the monkeys' classificatory behavior.

Given the nature of the errors described above, one might be tempted to hypothesize that the concept controlling the categorization behavior of the monkeys was animate/inanimate. This seems not to be the case. A horse appeared quite prominently in a new nonperson slide of Test 3 in Experiment 4. Only a single error was elicited by this slide, and it occurred on the fifth test session.

Concluding Comments

The conclusion that emerges from our results is that if the person concept was implicated in the monkeys' categorization behavior, it was not the sole, or perhaps not even the dominant, source of control. It is true that the monkeys' transfer performance to new exemplars was impressive, whether or not they contained a red patch. But having identified one controlling constellation of irrelevant features that resulted in errors to nonperson exemplars, one wonders whether correct responses to positive instances were not largely the product of limited constellations of *relevant* features formed by the same discriminative mechanisms. That is to say, rather than assembling from the relevant features in the various positive exemplars some sort of abstract representation of a person, the monkeys might have identified a limited number of features, alone or in combination, such as a pair of eyes contained within a closed oval, which, if present in a new slide to a reasonable degree of similarity, would elicit a "person" categorization (cf. Cerella's, 1982, template model). Although it is a form of complex discriminative behavior that deserves serious study in its own right, the egregious and persistent errors to which this process is susceptible would seem to limit its usefulness as a model of concept formation or identification.

After devoting a fair amount of effort to the present research issue, we have come to the view that continued efforts along similar lines are not likely to prove very fruitful. There are a number of reasons for this pessimistic conclusion. First, because there is no general agreement as to the critical processes by which natural category concepts are formed in animals or in humans—or even as to the definition of such concepts—the objective of the research is nebulous. Simply because an animal classifies slides in accordance with an experimenter-defined concept does not necessarily imply that the animal's behavior is controlled by the same concept. A point that cannot be stressed too strongly is that comparable terminal performances can be achieved by very different underlying mechanisms. Pigeons learn to match visual stimuli at least as efficiently as do monkeys, but apparently they do not arrive at that result by exactly the same means (D'Amato, Salmon, & Colombo, 1985; D'Amato, Salmon, Loukas, & Tomie, 1986).

It has been said in this connection that animals may not show a particular cognitive capacity unless specifically reinforced for doing so (e.g., Herrnstein, 1985). But the previous

caveat applies here as well. Young children, but neither monkeys nor pigeons, spontaneously (i.e., at the first opportunity) display associative symmetry (D'Amato, Salmon, Loukas, & Tomie, 1985; Sidman, Rauzin, Lazar, Cunningham, Tailby, & Carrigan, 1982). If trained on a series of forward and backward conditional matching tasks, monkeys and pigeons eventually might come to anticipate the backward task and respond appropriately. But such an outcome would not necessarily mean that the same cognitive capacity that allowed the children to spontaneously display associative symmetry—a capacity that probably contributes to humans' more flexible associative networks (cf. Asch & Ebenholtz, 1962)—was also operating in the animals. Additional tests, perhaps with different classes of stimuli (cf. D'Amato, Salmon, & Colombo, 1985), would be required to establish this parallel.

The problem, then, is that in order to be sure that successful categorization behavior in animals has a conceptual basis, we must rule out alternative possibilities, which, for both empirical and theoretical reasons, is generally a very difficult task.

Another reason for our disenchantment is that it is extremely difficult to analyze with any completeness the specific features and/or constellation of features that animals use in their classification of slides or similar representations of natural object concepts. The very strengths of such stimulus materials—complexity and variability—become their undoing when it comes to the important task of specifying the controlling stimuli. Less complex stimuli will doubtless render the latter task more manageable (Cerella, 1986).

A quite different concern is that experimental instantiations of natural category concepts are usually impoverished versions of the real thing. Anyone who has worked with monkeys for an extended period of time would have difficulty believing that their responses to humans are totally devoid of a conceptual basis. The nature of their reactions to people, familiar and unfamiliar, seems to require such a capacity. But real people walk, talk, have odors, put food in their mouths as monkeys do, reciprocally look monkeys in the eyes, and so forth. In short, monkeys could very well have a well-developed concept of humans that might not be revealed by two-dimensional, lifeless representations. The same possibility obviously exists for many of the other natural category concepts investigated in animals.

More promising, perhaps, is the search for basic mechanisms by which animals learn to differentiate among simple and complex stimulus displays. For example, pigeons that learned to discriminate between one versus two arbitrary forms drawn by the same person generalized the discrimination to new instances; however, they failed miserably when the forms were created by another hand (Greene, 1983). What were the common features of the first set of forms that set them apart from the second? More generally, what variables determine whether, and to what strength, an irrelevant feature will gain control over an animal's discriminatory behavior, what factors govern the correlation or compounding of individual features, and how do these basic discriminative processes vary from one species to another?

References

- Asch, S. E., & Ebenholtz, S. M. (1962). The principle of associative symmetry. *Proceedings of the American Philosophical Society*, 106,

- 135-163.
- Andrew, G., & Harlow, H. F. (1948). Performance of macaque monkeys on a test of the concept of generalized triangularity. *Comparative Psychology Monographs*, 19 (3, Serial No. 100).
- Cerella, J. (1979). Visual classes and categories in the pigeon. *Journal of Experimental Psychology: Human Perception and Performance*, 5, 68-77.
- Cerella, J. (1982). Mechanisms of concept formation in the pigeon. In D. J. Engle, M. A. Goodale, & R. J. W. Mansfield (Eds.), *Analysis of visual behavior* (pp. 241-260). Cambridge, MA: MIT Press.
- Cerella, J. (1986). Pigeons and perceptrons. *Pattern Recognition*, 19, 431-438.
- D'Amato, M. R., Salmon, D. P. (1983, August). *Person concept in cebus monkeys*. Paper presented at the meeting of the American Psychological Association, Anaheim, CA.
- D'Amato, M. R., Salmon, D. P., & Colombo, M. (1985). Extent and limits of the matching concept in monkeys (*Cebus apella*). *Journal of Experimental Psychology: Animal Behavior Processes*, 11, 35-51.
- D'Amato, M. R., Salmon, D. P., Loukas, E., & Tomie, A. (1985). Symmetry and transitivity of conditional relations in monkeys (*Cebus apella*) and pigeons (*Columba livia*). *Journal of Experimental Analysis of Behavior*, 44, 35-47.
- D'Amato, M. R., Salmon, D. P., Loukas, E., & Tomie, A. (1986). Processing of identity and conditional relations in monkeys (*Cebus apella*) and pigeons (*Columba livia*). *Animal Learning & Behavior*, 14, 365-373.
- Greene, S. L. (1983). Feature memorization in pigeon concept formation. In M. Commons, R. J. Herrnstein, & A. R. Wagner, (Eds.), *Quantitative analyses of behavior: Discrimination processes* (Vol. 4, pp. 209-229). Cambridge, MA: Ballinger.
- 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. (1985). Riddles of natural categorization. *Philosophical Transactions of the Royal Society of London*, B308, 129-144.
- Herrnstein, R. J., & Loveland, D. H. (1964). Complex visual concept in the pigeon. *Science*, 146, 549-551.
- Herrnstein, R. J., & de Villiers, P. A. (1980). Fish as a natural concept for people and pigeons. In G. H. Bower (Ed.), *The psychology of learning and motivation* (Vol. 14, pp. 59-95). New York: Academic Press.
- Lea, S. E. G. (1984). In what sense do pigeons learn concepts? In H. L. Roitblat, T. G. Bever, & H. S. Terrace (Eds.), *Animal cognition* (pp. 263-276). Hillsdale, NJ: Erlbaum.
- Lehr, E. (1967). Experimentelle Untersuchungen an Affen und Halbaffen über Generalisation von Insekten- und Blütenabbildungen [Experimental investigations of the generalization of insect and flower pictures by monkeys and a lemur]. *Zeitschrift für Tierpsychologie*, 24, 208-244.
- Medin, D. L., & Smith, E. E. (1984). Concepts and concept formation. *Annual Review of Psychology*, 35, 113-138.
- Ringo, J. L., & Doty, R. W. (1985). A macaque remembers pictures briefly viewed six months earlier. *Behavioural Brain Research*, 18, 289-294.
- Schrier, A. M., Angarella, R., & Povar, M. L. (1984). Studies of concept formation by stump-tailed monkeys: Concepts humans, monkeys, and letter A. *Journal of Experimental Psychology: Animal Behavior Processes*, 10, 564-584.
- 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.
- Shank, R. C., Collins, G. C., & Hunter, L. E. (1986). Transcending inductive category formation in learning. *Behavioral and Brain Sciences*, 9, 639-686.
- Sidman, M., Rauzin, R., Lazar, R., Cunningham, S., Tailby, W., & Carrigan, P. (1982). A search for symmetry in the conditional discriminations of rhesus monkeys, baboons, and children.
- Siegel, R. K., & Honig, W. K. (1970). Pigeon concept formation: Successive and simultaneous acquisition. *Journal of the Experimental Analysis of Behavior*, 13, 385-390.
- Vaughan, W., Jr., & Greene, S. L. (1984). Pigeon visual memory capacity. *Journal of Experimental Psychology: Animal Behavior Processes*, 10, 256-271.
- Yoshikubo, S. (1985). Species discrimination and concept formation by rhesus monkeys (*Macaca mulatta*). *Primates*, 26, 285-299.

Received April 14, 1987

Revision received July 21, 1987

Accepted July 23, 1987 ■