Mechanisms of *Same/Different* Abstract-Concept Learning by Rhesus Monkeys (*Macaca mulatta*)

Jeffrey S. Katz
Auburn University

Anthony A. Wright and Jocelyne Bachevalier
University of Texas Medical School at Houston

Experiments with 9 rhesus monkeys (*Macaca mulatta*) showed, for the first time, that abstract-concept learning varied with the training stimulus set size. In a *same/different* task, monkeys required to touch a top picture before choosing a bottom picture (*same*) or white rectangle (*different*) learned rapidly. Monkeys not required to touch the top picture or presented with the top picture for a fixed time learned slowly or not at all. No abstract-concept learning occurred after 8-item training but progressively improved with larger set sizes and was complete following 128-item training. A control monkey with a constant 8-item set ruled out repeated training and testing. Contrary to the unique-species account, it is argued that different species have quantitative, not qualitative, differences in abstract-concept learning.

Abstract concepts are the basis of much of our so-called higher order cognitive processing (e.g., language and mathematics). Abstract concepts are rules about relationships such as identity or sameness. Children apparently develop cognition in stages and expand their abstract concept of sameness to include number length, area, and volume (Piaget & Inhelder, 1966/1969). Indeed, William James (1890/1950) declared that our "sense of sameness is the very keel and backbone of our thinking" (p. 459). In the laboratory, the abstract concept of sameness is usually studied in matching-to-sample (MTS) or *same/different* (S/D) tasks. In MTS, subjects typically view a sample stimulus and then choose one of two comparison stimuli, with the correct choice being the comparison that matches the sample. In S/D, subjects view a pair of stimuli and then make one of two responses to indicate whether the stimuli are the *same* or *different*. The determination of abstract-concept learning in both cases is accurate performance with novel test stimuli. Accurate performance with novel stimuli means that the subject has learned an abstract rule that transcends the particular training stimuli. Such transfer performance is what makes abstract-concept learning unique and different from other forms of concept learning.

Abstract concepts are different from what are called "natural" concepts. Natural concepts are categories of items bound by absolute features of objects such as cars, chairs, flowers, person, water, or trees (e.g., Bhatt, Wasserman, Reynolds, & Knauss, 1988; Herrnstein, Loveland, & Cable, 1976; Roberts & Mazma-nian, 1988). Natural concepts involve learning specific features shared by a class or category of items (e.g., prototype). By contrast, abstract concepts do not involve learning specific stimulus features. Instead, they involve learning the relationships between items. Thus, abstract concepts involve relational learning as opposed to the item-specific learning of natural-concept learning. Nothing more in this article will be said about natural-concept learning. This article deals exclusively with abstract-concept learning and will focus mainly on S/D abstract-concept learning by rhesus monkeys.

A great deal of interest in S/D abstract-concept learning was generated by a seminal article in 1978 entitled "On the Abstractness of Human Concepts: Why It Would Be Difficult to Talk to a Pigeon" by David Premack. Among the hypotheses Premack advanced in that article (also in Premack, 1983; Premack & Premack, 1983) was the claim that nonlanguage trained organisms were incapable of learning S/D abstract concepts. Premack claimed that the ability to learn S/D concepts reflected abstract thinking acquired through language training. Thus, S/D abstract-concept learning was, according to his hypothesis, limited to species that could learn language. Since the advancement of this hypothesis, language-naive chimpanzees and a variety of other species including baboons, rhesus monkeys, capuchin monkeys, parrots, and pigeons have been able to learn (wholly or partially) S/D abstract concepts (Bhatt & Wright, 1992; Bovet & Vauclair, 2001; Cook, Cavoto, & Cavoto, 1995; Pepperberg, 1987; Santiago & Wright, 1984; Thompson, Oden, & Boysen, 1997; Wasserman, Hugart, & Kirkpatrick-Steger, 1995; Wright, Santiago, & Sands, 1984).

These advancements of showing that many different species do have the cognitive capability to learn S/D abstract concepts were made possible by new procedures to train and test abstract-concept learning. These procedures better capitalized on the predispositions of the species being tested and thereby increased the chances of revealing abstract-concept learning. Notwithstanding their important achievements in showing the widespread ability of different species to learn S/D abstract concepts, these experiments have been more or less one-shot demonstrations. Procedures were tried and either the subjects succeeded or failed to learn the abstract

Correspondence concerning this article should be addressed to Jeffrey S. Katz, Department of Psychology, Auburn University, 226 Thach Hall, Auburn, Alabama 36849, or Anthony A. Wright or Jocelyne Bachevalier, Department of Neurobiology and Anatomy, University of Texas Medical School at Houston, P.O. Box 20708, Houston, Texas 77225. E-mail: katzjef@auburn.edu, anthony.a.wright@uth.tmc.edu, or jocelyne.bachevalier@uth.tmc.edu

concept. There has been little or no attempt to understand which aspects of the procedure led to abstract-concept learning and which did not. With all of the failures and successes of S/D and MTS abstract-concept learning over the years, we believe it is time to shift the focus from which species do and do not have the cognitive capabilities to learn abstract concepts to the processes and mechanisms by which concepts are learned. This shift involves identifying and exploring aspects of the task that are critical for abstract-concept learning. *Critical* means that over the range of the parameter (e.g., training stimulus set size), concept learning will vary from little or no concept learning to very accurate, possibly complete, concept learning. Discovery of such critical parameters and how they affect abstract-concept learning, we believe, will provide the necessary basis for understanding mechanisms and processes of how abstract concepts are learned. The studies reported in this article provide a beginning of this identification and exploration of critical parameters in S/D abstract-concept learning by rhesus monkeys with video picture stimuli and a touch screen.

Related to which aspects of the task are critical for abstract-concept learning is what constitutes abstract-concept learning. Our goal in the experiments of this article was to obtain full (complete) concept learning. We refer to concept learning as being "full" when novel-stimulus test performance is equivalent to baseline training performance and both are of high accuracy (e.g., $\geq 80\%$ correct). If novel-stimulus test performance is intermediate, that is, above chance (50% correct) but less than training performance, then the best that can be said for such a result is that partial concept learning was obtained. More to the point, we believe that partial concept learning should only be a beginning and that by identifying and manipulating critical parameters full concept learning should be obtainable.

In this article, we identify two parameters, multiple observing responses and training stimulus set size, that are critical to S/D abstract-concept learning by rhesus monkeys. When we began these studies, it was unclear whether abstract-concept learning generally depends on training with a large number of stimuli or exemplars. One could point to the successful MTS abstract-concept learning with monkeys and pigeons (e.g., Moon & Harlow, 1955; Overman & Doty, 1980; Weinstein, 1941; Wright, Cook, Rivera, Sands, & Delius, 1988) with large (e.g., 100+ items) stimulus sets. Comparably large stimulus sets have also been used to train monkeys to learn S/D abstract concepts (e.g., Bhatt & Wright, 1992; Wright et al., 1984; Wright, Shyan, & Jitsumori, 1990). Supporting this evidence for large set sizes and abstract-concept learning are failures of MTS abstract-concept learning with small stimulus sets (e.g., Berryman, Cumming, Cohen, & Johnson, 1965; Cumming & Berryman, 1961; Cumming, Berryman, & Cohen, 1965; D'Amato, Salmon, & Colombo, 1985; Farthing & Opuda, 1974; Holmes, 1979). A set-size effect on concept learning makes perfect sense. With small stimulus sets, individual features and patterns of the displays might become the controlling cue, whereas with large stimulus sets individual stimulus features change so often that stimulus relationships (e.g., S/D) become the controlling cue. Nevertheless, despite this evidence and common-sense appeal, there are examples of partial and even full MTS and S/D abstract-concept learning with very small to moderate stimulus set sizes with dolphins, sea lions, chimpanzees, orangutans, and (even) pigeons (e.g., Herman, Hovancik, Gory, & Bradshaw, 1989; Kastak & Schusterman, 1994; Oden, Thompson, & Premack, 1988; Robinson, 1955; Wright, 1997). From this

mixed evidence, it is difficult to extract any general principle with regards to set size by comparing across different experiments. To determine the role of set size in abstract-concept learning, set size needs to be manipulated within the same experiment, preferably within the same subject (cf. Young, Wasserman, & Dalrymple, 1997). If subjects did not learn the concept with a small stimulus set size but then were to increase their concept learning with increasing set size, this would be very strong evidence for the functional and critical role of set size in abstract-concept learning. This was the plan of the experiments reported in this article.

Our plan was to train rhesus monkeys with a small stimulus set of 8 items, test for concept learning, expand the set size, and repeat the process several times. However, the first group of monkeys (Experiment 1, 0-response group) that we "trained" in the task, for the most part, did not learn. Learning the task was essential to continuing the experiments. Therefore, we trained a second group (Experiment 1, 10-response group) of monkeys to touch an initial stimulus 10 times (observing-response requirement) before being presented with the second stimulus. This observing-response requirement promoted rapid learning of the task. There was no concept learning with this 8-item set, which was a good result because it allowed us to test whether increases in set size might lead to abstract-concept learning. We expanded the set size and tested for abstract-concept learning (Experiment 4), after controlling for exposure (as opposed to responses) to the first stimulus (Experiment 2), and after retraining monkeys that had not learned the task from the first two experiments on the 10-response procedure (Experiment 3).

Experiment 1

Experiment 1 tested whether responding to a first-presented stimulus of a pair would hasten learning relative to a group not making such responses. Some evidence suggested that contact with the stimulus, like touching the pictures, should improve the rate of learning (e.g., Harrison, Iverson, & Pratt, 1977; Meyer, Treichler, & Meyer, 1965; Stollnitz, 1965). Other evidence indicated that stimulus contact might not be necessary. Monkeys learned an S/D task (and concept) by moving a lever (e.g., Sands & Wright, 1980a, 1980b; Wright et al., 1984) and learned a slightly different task by moving a joystick (e.g., Rumbaugh, Richardson, Washburn, Savage-Rumbaugh, & Hopkins, 1989; Washburn, Hopkins, & Rumbaugh, 1989). Like the issue on set size, other aspects of the procedure undoubtedly contribute to ease or difficulty in learning and may obscure the effect of stimulus contact. To determine the role of stimulus contact in learning, stimulus contact needs to be manipulated within the same experiment. This was the purpose of Experiment 1.

Method

Animals

One female and 6 male rhesus monkeys (*Macaca mulatta*) participated in this experiment. At the beginning of the experiment, the ages of the monkeys were 3–8 years old. There were two groups of monkeys. A 0-response group consisted of Gracie, P2, and Slim; these monkeys were experimentally naive. A 10-response group consisted of Boobah, Cuba, Scuba, and Cubby. Boobah, Cuba, and Scuba had a prior history of listening to monkey calls in a different experimental chamber at another institution but no instrumental training; Cubby was experimentally naive.

All monkeys were individually housed in cages located in a room shared with other rhesus monkeys. Experimental sessions were conducted 5–7 days a week. On testing days, access to food (Purina Monkey Chow) and water in their home cages was restricted about 15 hr before testing. On nontesting days, access to food and water was unrestricted. Additionally, vegetable and fruit supplements were provided at the end of each week. A 12-hr light–dark cycle was maintained in the colony.

Apparatus

Chamber. Monkeys were tested unrestrained in one of two identical (47.50 cm wide × 53.13 cm deep × 66.25 cm high) custom aluminum test chambers (see Bhatt & Wright, 1992, for further details). A fan (Dayton 4C440, Niles, IL), located in the ceiling of the chamber, provided ventilation and white noise. Bio-Serv banana pellets (300 mg) were dispensed (Model ENV-203-300, MED Associates, Inc., Georgia, VT) into a pellet cup (5.60 cm diameter and 2.50 cm deep) that was 10.00 cm from the left edge and 52.50 cm from the top edge of the front panel. Tang orange drink was delivered (Liquid solenoid valve, Model 71215, Honeywell, Inc., New Britain, CT) through a juice spout that was 10.00 cm from the right edge and 42.50 cm from the top edge of the front panel. Touches to a computer monitor were detected by an infrared touch screen bezel (Model 81009703-01, Carroll Touch, Round Rock, TX). The bezel fit snugly within a 40.00 × 33.75-cm cutout in the front panel that was centered 9.38 cm from the top of the operant chamber. Touch responses were directed by a Plexiglas template (32.50 × 40.00 cm) with cutouts matching the size and location of the stimuli.

Stimuli. Travel-slide color pictures were digitized using a Howtek (Hudson, NH) Photomaster (87RU) camera and a TrueVision (Indianapolis, IN) TARGA-16 processing card in a 256 × 256 resolution. Stimuli were presented on a 39.00-cm color monitor (Eizo FX-C6, 600 × 480-pixel resolution, Ishikawa, Japan). Stimulus displays consisted of two travel-slide color pictures (each 13.75 × 9.70 cm) and a white rectangle (6.25 × 5.60 cm) on a black background. The pictures were vertically aligned with a 3.40-cm gap between them. The top picture was centered 20.63 cm from the left edge and 18.75 cm from the top of the front panel. The bottom of the white rectangle was horizontally aligned with the bottom of the lower picture with a 3.70-cm gap between them. S/D trials were composed by quasirandomly selecting pictures from an 8-item set (apples, buildings, cat, face, flower, glass and pitcher, keys, and orangutan; see Figure 1). There were a total of 64 different displays (8 *same* and 56 *different*).

Experimental control. Experimental events were controlled and recorded using custom software written in Visual Basic on a Pentium personal computer. A video card (Jeronimo J2 Advanced Graphics Accelerator, Appian Graphics, Redmond, WA) controlled the monitor. A computer-controlled relay interface (Model PI0–12, Metrabyte, Taunton, MA) operated the liquid solenoid valve and pellet dispenser.

Procedure

Preliminary training. Monkeys were first trained to eat banana pellets (300 mg; Bio-Serve, Frenchtown, NJ) from the food cup and drink Tang orange juice from the juice spout. Responses to the video monitor and touch screen were shaped by successive approximations to the white rectangle (later to become the *different* response area) and another white rectangle (13.75 × 9.70 cm) placed in the lower picture position. These two rectangles were presented on separate trials and randomly occurred equally often in a 100-trial session. A single touch to the white rectangle was followed by a 1.1 s, 660 Hz tone. The reinforcer type, food (1 pellet) or juice (3–5 cc depending on the monkey), randomly occurred after a touch response with the constraint that an equal number of pellet and juice rewards followed both responses. Food reward was delivered simultaneously with the tone. Juice reward was delivered 1 s after the tone. A 15-s intertrial interval (ITI) followed reinforcement. Once a monkey was consistently responding (1–5 sessions), S/D training began.

Same/different training. For the monkeys of the 0-response group (Gracie, P2, and Slim), a trial began with presentation of the two pictures and the white rectangle (see Figure 1). If the two pictures were the *same*, a touch response to the lower picture was rewarded. If the two pictures were *different*, a touch response to the white rectangle was rewarded. For the monkeys of the 10-response group (Boobah, Cuba, Cubby, and Scuba), a trial began with presentation of the upper picture. Initially, these monkeys were required to touch the upper picture only once. This (observing) response was followed by presentation of the lower picture and the white rectangle along with the upper picture. Thus, the final displays for both groups were identical. The monkeys then made a choice response, like those of the 0-response group. The number of observing responses required for the 10-response group was increased by one response following sessions completed in less than 90 min. The monkeys reached the 10-response (FR-10) requirement by the 20th session.

Following a choice, the display was turned off. A correct choice resulted in tone and reward. The percents of juice and pellet rewards varied slightly according to the individual monkey's preference but were equated over *same* and *different* trial types. An incorrect choice resulted in no reward. A correct or incorrect choice was followed by a 15-s ITI. Starting on the fifth training session, incorrect choices were also followed by a repeat of the incorrect trial (correction procedure). On correction trials, a 15-s timeout preceded the ITI. Accuracy was based on first trial performance only. Performance on correction trials (although recorded) does not figure in any analyses presented in this article.

Sessions consisted of 100 trials (50 *same*/50 *different*). If an animal did not complete a session, then the session was continued on the next day. This occasionally happened at the beginning of training. Training continued until performance was 80% or better on three consecutive sessions. The correction procedure was then removed and training continued until the same criterion was met.

Results and Discussion

Figure 2 shows percent correct accuracy across sessions for both groups. Results from monkeys in the 10-response group are shown by unfilled symbols, and results from the 0-response group are shown by filled symbols. Monkeys in the 10-response group rapidly learned the task in an average of 4,050 trials. By contrast, 2 of the monkeys in the 0-response group did not learn the task even after 25,000 training trials. One monkey, P2, from the 0-response group did learn the task but learned much more slowly than any monkey from the 10-response group.

The results from this experiment clearly show that learning benefits from requiring multiple observing responses to the first presented item of the S/D task. These observing responses require contact with the stimuli and have been shown to facilitate learning in other discrimination tasks (e.g., Harrison et al., 1977; Meyer et al., 1965; Stollnitz, 1965). The present study extends this work by showing the importance of multiple stimulus contacts. Even the 0-response group had to contact the stimuli when they made their choice responses, but they did not learn, for the most part.

Multiple observing responses may help focus the subjects' attention on the elements that make up each stimulus pair. Such attention might then increase the chances that the subject would notice that the second element is the *same* as the first element or *different* from it. If so, the function of observing responses in the present study might be related to observing responses of pigeons learning an MTS abstract concept (Wright, 1997). Pigeons that pecked a sample 20 times before being presented with two (choice) stimuli learned the task more rapidly than those that did not peck the sample (Wright, 2001). Moreover, pigeons in the 20-response group learned the MTS concept, whereas those in the 0-response

Set of Eight Training Pictures

10-Response Group
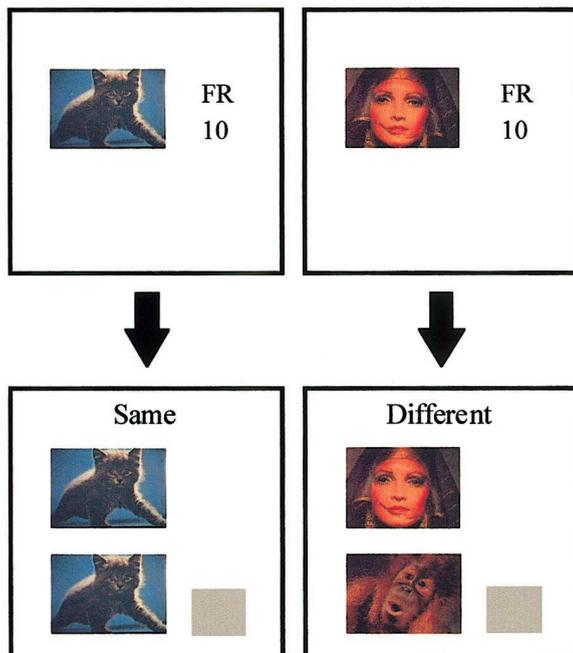
0-Response Group
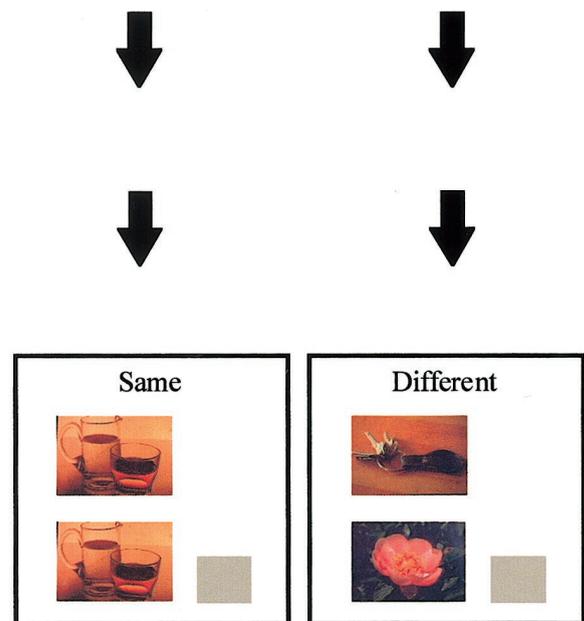
Figure 1. Eight pictures used in the training of the *same/different* task in Experiments 1–3 (top panel). The 10-response (bottom-left panel) and 0-response (bottom-right panel) procedures used in Experiments 1, 3, and 4. The 10-response procedure required monkeys to touch the top picture 10 times (FR 10) before receiving the lower picture and gray rectangle and allowing a choice response. In both procedures, a touch to the bottom picture was correct on *same* trials. A touch to the gray rectangle was correct on *different* trials. Examples of *same* and *different* trials are proportional to the actual displays, but actual displays were color pictures on a black background with a white rectangle.

group did not. Similarities to this pigeon concept-learning task may end here because, as we shall see in Experiment 4, none of the monkeys learned the concept following training with the 8-item set. Before presenting those results and pursuing the study of set size on concept learning, we will present two additional experiments to help establish and define the role of observing responses in S/D acquisition.

Experiment 2

Monkeys in the 10-response group learned the S/D task much more rapidly than any monkey in the 0-response group. Although multiple responses to the first item appear to be the critical differ-

ence, this item was presented for several seconds before the pair of items was presented together. One could argue that the initial presentation of the upper item might have been important for learning the task, not responses to the item. To test this possibility, the 2 monkeys (Gracie and Slim) that did not learn the task were retrained with the upper item presented alone prior to presenting the pair of pictures together. The length of time that the upper item by itself was presented was made equivalent to the average time required for Boobah, Cuba, and Scuba of the 10-response group to complete their response requirement. Because these monkeys' previous experience in Experiment 1 might affect learning in Experiment 2, an experimentally naive monkey, Minnie, was added to this group.
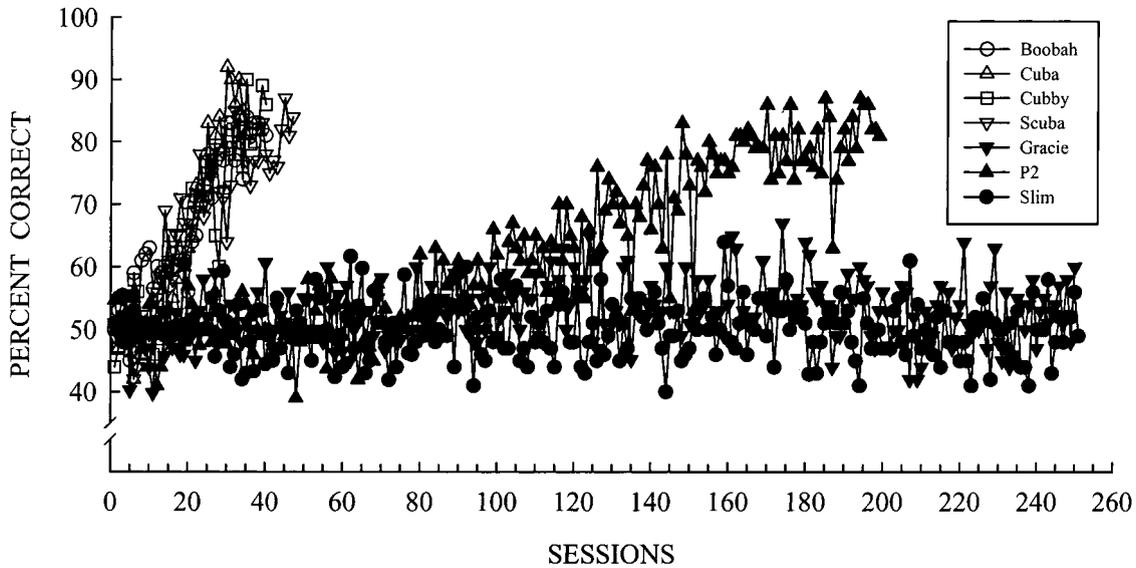
Figure 2. Percent correct across 100-trial sessions from Experiment 1 for the 10-response group (unfilled symbols) and the 0-response group (filled symbols).

Method

Animals

Gracie and Slim from Experiment 1 and a 6-year-old experimentally naive female rhesus monkey, Minnie, participated in the experiment. Minnie was given the same preliminary training that the other monkeys had received before beginning Experiment 1.

Apparatus and Stimuli

The same apparatus and the same 8 pictures from Experiment 1 were used.

Procedure

All aspects of the experiment were the same as in Experiment 1, except that the upper item was presented for a fixed time. The exposure duration for the upper item was set to the average time for Boobah, Cuba, and Scuba from Experiment 1 to complete their response requirement. This fixed time increased over the first 20 sessions (when the fixed ratio was increased from 1 to 10 for the 10-response group) and then was fixed at 10 s thereafter. Observations of the monkeys with a video camera revealed that, for the most part, they observed the upper item during the fixed-time display. Following the fixed-time period, the lower item and the white rectangle were simultaneously presented along with the upper picture. For any of the monkeys in this experiment, training was stopped after 30 sessions if there were no clear indications of acquisition.

Results and Discussion

S/D performances of the individual monkeys in Experiment 2 are shown in Figure 3 as filled symbols. For comparison, the mean acquisition results of monkeys in the 10-response group of Experiment 1 are shown as unfilled symbols. Minnie and Slim showed no signs of acquisition after 30 sessions of training. Therefore, their training in this experiment was stopped. Gracie did learn the task. Her training was continued until she met the criterion of two noncorrection sessions with performance better than 80%. Gracie

learned the task at about the same rate as the 10-response group (Boobah, Cuba, Scuba, and Cubby).

A further analysis of Gracie's performance shows that she spontaneously touched the items during the presentation of the upper item alone. Although Gracie was not required to touch the upper item during its presentation, she did touch the upper item on 26% of the trials. Furthermore, an analysis of the last 32 acquisition sessions (where learning was greatest) showed that Gracie was more accurate on trials where she touched the upper item (82%) than on those trials where she did not touch (71%), $t(31) = 6.99$, $p < .001$. Minnie also occasionally touched the upper item but did so on only 12% of the trials and was not significantly more accurate on "touched" trials (51%) than on "nontouched" trials
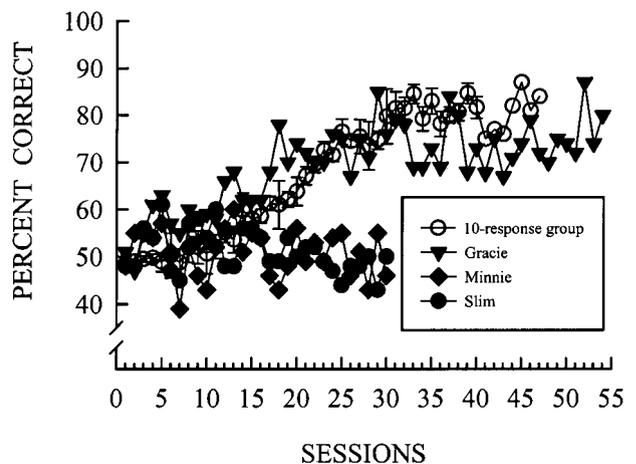
Figure 3. Percent correct across 100-trial sessions from Experiment 2 for 3 monkeys trained on the fixed-time procedure (filled symbols) compared with the average of the 10-response group from Experiment 1 (unfilled symbols).

(43%), $t(29) = 1.53$, $p > .05$. Slim showed no learning of the task, and he virtually never touched the upper item.

In summary, Minnie and Slim failed to learn the task with the fixed-time procedure. Gracie touched during the fixed presentation time, which may have been instrumental in her learning the task. Thus, regardless of whether it is required, touching the upper item (cf. 10-response group, Experiment 1) may be an important component in learning this S/D task, not simply exposure to the upper item for a period of time.

Experiment 3

Slim and Minnie did not learn with the fixed-time procedure of Experiment 2. Therefore, an additional test of the importance of observing responses was to train these monkeys on the 10-response procedure. If they learned the task, there would then be additional evidence that observing responses were critical in learning this S/D task. This objective was the purpose of Experiment 3. A third monkey, Aruba, who, unlike Minnie and Slim, did not have any fixed-time training, was added to this group.

Method

Animals

Minnie and Slim from Experiment 2 and a female rhesus monkey (Aruba, 7 years old) participated in the experiment. Aruba had 86 sessions of training in the 0-response procedure of Experiment 1 before beginning training in this experiment.

Apparatus, Stimuli, and Procedure

The same apparatus, 8 pictures, and procedures, including the 10-response procedure of Experiment 1, were used to train the monkeys in this experiment. Because they all had previous experience in the task, the preliminary training procedures of Experiment 1 were unnecessary.

Results and Discussion

The acquisition results for the individual monkeys from Experiment 3 are shown in Figure 4 as filled symbols. Mean results for the 10-response group from Experiment 1 are shown as unfilled symbols for comparison. All 3 monkeys rapidly learned the S/D task with the 10-response procedure. Indeed, they acquired the task somewhat more rapidly than the mean of the 10-response group of Experiment 1 as confirmed by a significant interaction from a two-way mixed analysis of variance (ANOVA) of 3-Session Blocks × Group, with blocks serving as the repeated measure, $F(8, 40) = 2.75$, $p < .017$. This result is doubly important because it shows not only the importance of observing responses but also that the previous training of these monkeys (more than a year in the case of Slim) did not adversely effect learning relative to an experimentally naive group. The slight advantage of these monkeys over the experimentally naive group may have been due to their prior familiarity with the stimuli and environment despite not learning the task.

In summary, Experiments 1, 2, and 3 combine to show that making observing responses (i.e., touching) to the first presented item is usually necessary for learning this S/D task. It is important to emphasize that the critical role of observing responses may be particular to this S/D task with video monitor, touch screens, and rhesus monkeys. Rhesus monkeys that moved a lever to the right
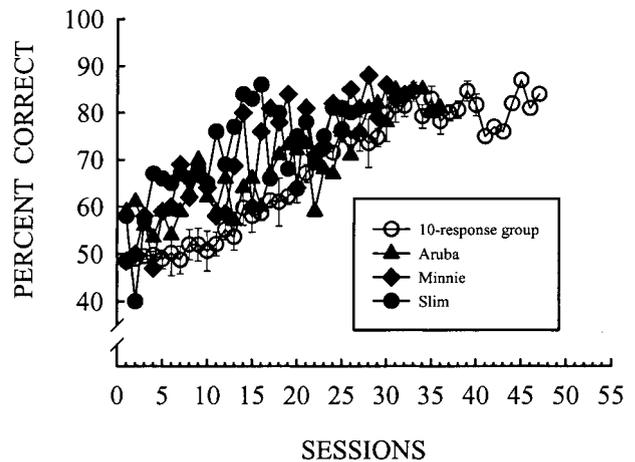
Figure 4. Percent correct across 100-trial sessions from Experiment 3 for monkeys trained on the 10-response procedure (filled symbols) after failing to acquire the discrimination on the 0-response or fixed-time procedures. Also shown is the average of the 10-response group from Experiment 1 (unfilled symbols).

or left to indicate *same* or *different*, respectively, learned the task (e.g., Sands & Wright, 1980a, 1980b; Wright et al., 1984). Although these monkeys did not make observing responses, they did start trials by a downward press of the lever. Even if observing responses are not absolutely critical with other species, tasks, or physical set-ups, they appear to produce more rapid learning. Clearly, the S/D task is not an easy task, and therefore learning differences might be more apparent than in somewhat easier tasks. The S/D task, and in particular the abstract S/D concept, has been thought to be something that only the most intelligent species with language training can learn (Premack, 1978, 1983). Whether these monkeys of Experiments 1–3 learned the S/D concept and, if not, what it takes for them to learn the concept is the topic of the next experiment.

Experiment 4

After learning the S/D task, the monkeys were tested for abstract-concept learning. We used conservative procedures to test and evaluate abstract-concept learning. We tested abstract-concept learning by intermixing only 10 novel-stimulus test trials with 90 baseline trials each session to avoid any possible disruption of overall performance in the task. Novel stimuli were used on all tests because any repetition of test stimuli has the potential to confound the interpretation. Potential confounds cannot be resolved by reinforcing either response, responses randomly, or no responses (extinction) because there is always some outcome and hence learning (i.e., confound) whenever a stimulus is presented. Our evaluation of abstract-concept learning was also conservative, and as previously mentioned, we consider that the abstract concept has been learned when novel-stimulus transfer performance is equivalent to baseline performance.

The monkeys trained in Experiments 1–3 were tested for abstract-concept learning. If they had not fully learned the concept, then we doubled the training stimulus set, retrained and retested the monkeys. This process of doubling of the training set and retraining and retesting was actually repeated four times. The

result of most interest was whether abstract-concept learning would increase with set size. If it did, particularly if full concept learning was eventually attained, then this result would be strong evidence for the instrumental role of set size in abstract-concept learning. One of the monkeys in this study, Cubby, was actually a control monkey for repeated training and testing experienced by the other monkeys. Cubby was maintained on the original 8-item training set but otherwise trained and tested like the other monkeys in the study.

Method

Animals and Apparatus

All 9 rhesus monkeys participated in the experiment. The same apparatus was used.

Procedure

On completion of learning the S/D task with the 8-item set (Experiment 1, 2, or 3; see Table 1 for a summary), transfer testing began the next session. Each testing session contained 100 trials (90 baseline and 10 transfer). Transfer testing was conducted over six consecutive sessions. Ten novel-stimulus test trials, 5 *same* and 5 *different*, were presented each test session. By definition of novelty, each picture was used only once during transfer testing. The 10 transfer trials were pseudorandomly placed within each test session, but none were placed within the first seven trials. There were 45 *same* and 45 *different* baseline trials each test session. Performance on transfer trials was rewarded in the same way that it was on baseline trials.

For the "experimental" monkeys (all except Cubby), the set size was increased from 8 to 16 pictures after six consecutive transfer test sessions. Set size was further increased to 32, 64, and 128 pictures, with transfer tests following criterion performance at these set sizes (but not the 16-item set). Each set retained the items from the previous set, and new pictures were added that had been used in prior transfer tests. Criterion performance with the 16-item set was 85% correct for three sessions with a correction procedure and one session of 85% correct after the correction procedure was removed. There was no correction procedure during training with 32-, 64-, and 128-item sets. Criterion performance with 32-, 64-, and 128-item sets was 85% correct for one session.

For the "control" monkey, Cubby, the picture set size remained fixed at 8 items. The number of training sessions (100 trials each session) for Cubby between transfer tests was the average required by other monkeys of the 10-response group (Boobah, Cuba, and Scuba) to reach criterion at each set size. Thus, Cubby was trained for 6, 1, and 1 session(s) prior to the second, third, and fourth transfer tests, respectively. Other aspects of the procedure were the same.

Table 1
Number of 100-Trial Sessions in Three Training Procedures Prior to Eight-Item Set Learning

Monkey	0-response procedure	Fixed-time procedure	10-response procedure
Gracie	250	54	0
P2	199	0	0
Slim	250	30	26
Minnie	0	30	32
Aruba	86	0	36
Boobah	0	0	40
Cuba	0	0	35
Cubby	0	0	40
Scuba	0	0	47

The observing-response requirements for the monkeys during successive cycles of training and transfer testing were the same as when the monkeys learned the task in either Experiment 1, 2, or 3. Thus, Aruba, Boobah, Cuba, Cubby, Minnie, Scuba, and Slim were required to touch the upper item 10 times—10-response procedure. P2 had both items presented simultaneously—0-response procedure. And finally, Gracie had the upper item presented for 10 seconds with no responses required—fixed-time procedure.

Results and Discussion

Figure 5 shows the number of trials (log scale) to criterion for each monkey at each set size. The dashed line shows the minimum number of trials according to the criterion at each set size. The functions show that for each monkey, the number of trials to criterion decreased as set size increased. This decrease in trial number was supported by a one-way repeated measures ANOVA, $F(4, 28) = 9.10$, $p < .001$.

Minnie, Boobah, Cuba, and Scuba, shown in the bottom row of Figure 5, learned close to the minimum number of trials for set sizes of 16, 32, 64, and 128 items. For the three largest sets, these monkeys learned in the minimum of one session (100 trials) in 11 out of 12 cases. It may be of some importance that across monkeys, there was a direct relation between the amount of no-response training (0-response or fixed-time) and trials to criterion. For example, Boobah, Cuba, and Scuba were trained only on the 10-response procedure and thus had no no-response training; they learned most rapidly. Minnie, whose trials-to-criterion function was similar to that for Boobah, Cuba, and Scuba, had only 3,000 trials of no-response (fixed-time) training prior to 10-response training. Aruba had 8,600 trials of 0-response training (prior to 10-response training) and required more trials to reach criterion than Minnie. Slim had 28,000 trials of 0-response training (prior to 10-response training) and required more trials to reach criterion than Aruba. (Gracie and P2 cannot be compared because they were not trained on the 10-response procedure.) Despite differences in trials to criterion across monkeys, all monkeys showed decreases in the number of trials to criterion for progressively larger set sizes. This decrease was substantial in every case and indicates a savings in training with new stimuli. Because the number of new items was in each case equal to the number of old items, it is difficult to come to any other conclusion but that the savings in training was due to concept learning. Nevertheless, tests of transfer to novel stimuli address this issue directly.

Figure 6 shows transfer performance following training at four of the set sizes. For each monkey, percent correct transfer performance (open circles) and baseline performance (filled circles) is averaged over the six transfer sessions following training at each set size. Performance on transfer trials increased with increasing set size, whereas baseline performance remained at a high level. This result was confirmed by a two-way repeated measures ANOVA of Set Size (8, 32, 64, 128) × Trial Type (baseline, transfer) on percent correct, which resulted in a significant interaction, $F(3, 21) = 17.80$, $p < .001$. There was a 30% mean increase in transfer performance across the set sizes (57%, 71%, 82%, 87%, for set sizes 8, 32, 64, and 128, respectively). This same pattern of increasing transfer was evident for all monkeys. Transfer performance initially (57%) was near chance (50%), indicating that whatever the monkeys learned depended on the particular training items and the resulting stimulus displays. Nevertheless, after four doublings of the set size, transfer performance
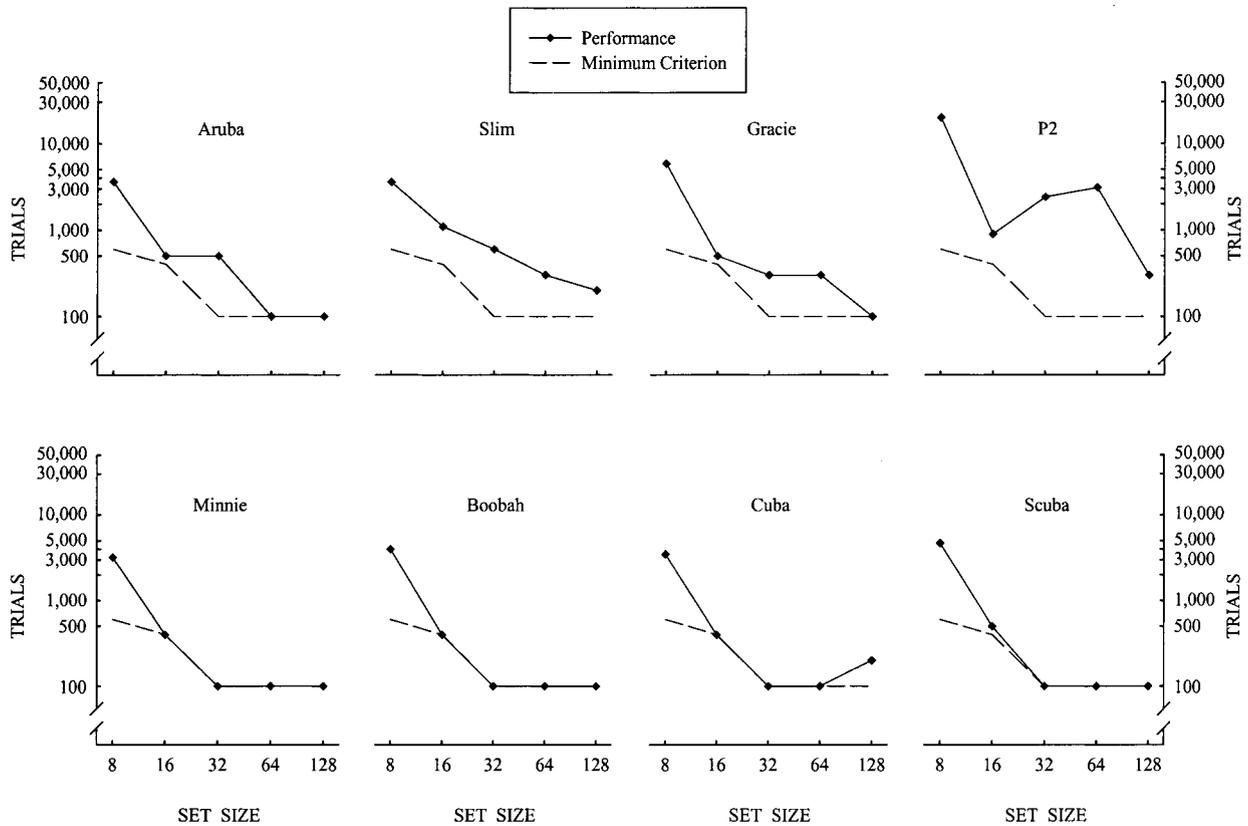
Figure 5. The number of trials to criterion (log scale) for each monkey (filled symbols) at each set size (8, 16, 32, 64, and 128 items). Dashed lines are the minimum trials to criterion at each set size.

(87%) with the 128 set size for 7 of the 8 monkeys was equivalent to baseline (89%), $t(7) = 1.04$, $p > .05$. Further analysis of transfer performance indicated that performance was stable across the six consecutive testing sessions at each set size, all $Fs(5, 35) < 1.00$, $p > .45$, and performance on *same* and *different* trials did not differ, all $Fs(1, 7) < 1.00$, $p > .69$. This result of transfer being equivalent to baseline means that after expanding the training set to 128 items, these monkeys had fully learned the abstract concept of *same/different*. The only exception was the partial concept learning by P2; P2's baseline remained at a high level (>80% correct), but transfer only rose to 68% correct.

The mean baseline and transfer results for Cubby (control) during the four transfer tests are shown in Figure 7. Also shown for comparison is the mean performance of Boobah, Cuba, and Scuba, the other monkeys from the 10-response group of Experiment 1. Cubby was trained and tested similar to the other monkeys of this group except that his set size was not expanded. The results generally support the conclusion that the set-size effect in the experimental monkeys was not merely an artifact of continued training and testing. There were, however, interesting trends in Cubby's transfer performance. Transfer initially increased with repeated transfer testing, but then, surprisingly, it decreased. Baseline performance remained at a high constant level. These differences between baseline and transfer were confirmed by a two-way repeated measures ANOVA of Testing Phase (1, 2, 3, 4) × Trial Type (baseline, transfer) on percent correct using sessions as the repeated measure, which resulted in a significant interaction, $F(3,$

15) $= 5.00$, $p < .05$. A trend analysis of the transfer data showed a significant quadratic component, $F(1, 5) = 11.00$, $p < .05$, substantiating that the function first rose (between the first and second test) and then fell (between the second, third, and fourth tests). Paired comparisons between transfer performance and chance (50%) showed that transfer performance significantly differed from chance on the second and third tests, $ts(5) > 5.90$, $p < .01$, but was not significantly different from chance on the final test, $t(5) = 0.81$, $p > .40$. This pattern of results is most curious. Because the final test performance was not different from chance, one has to conclude that there was no concept learning. But this conclusion does not take into account the considerable (75%, $SEM = 3.4$) transfer on the second test. This intermediate rise in transfer performance could not have resulted simply from repeated training and testing; otherwise it should have continued to rise. Instead, it fell over the latter half of this training and testing. One possibility is that this is the amount of training where the abstract concept could be most rapidly learned if new exemplars had been added to the training set. Independent of this possibility, Cubby's results are evidence that the set size effect shown by the other monkeys did not result from repeated training and testing.

General Discussion

Experiment 1 showed that responding 10 times to the first item (10-response group) before being presented a second item of a pair greatly enhanced learning of the S/D task relative to both items
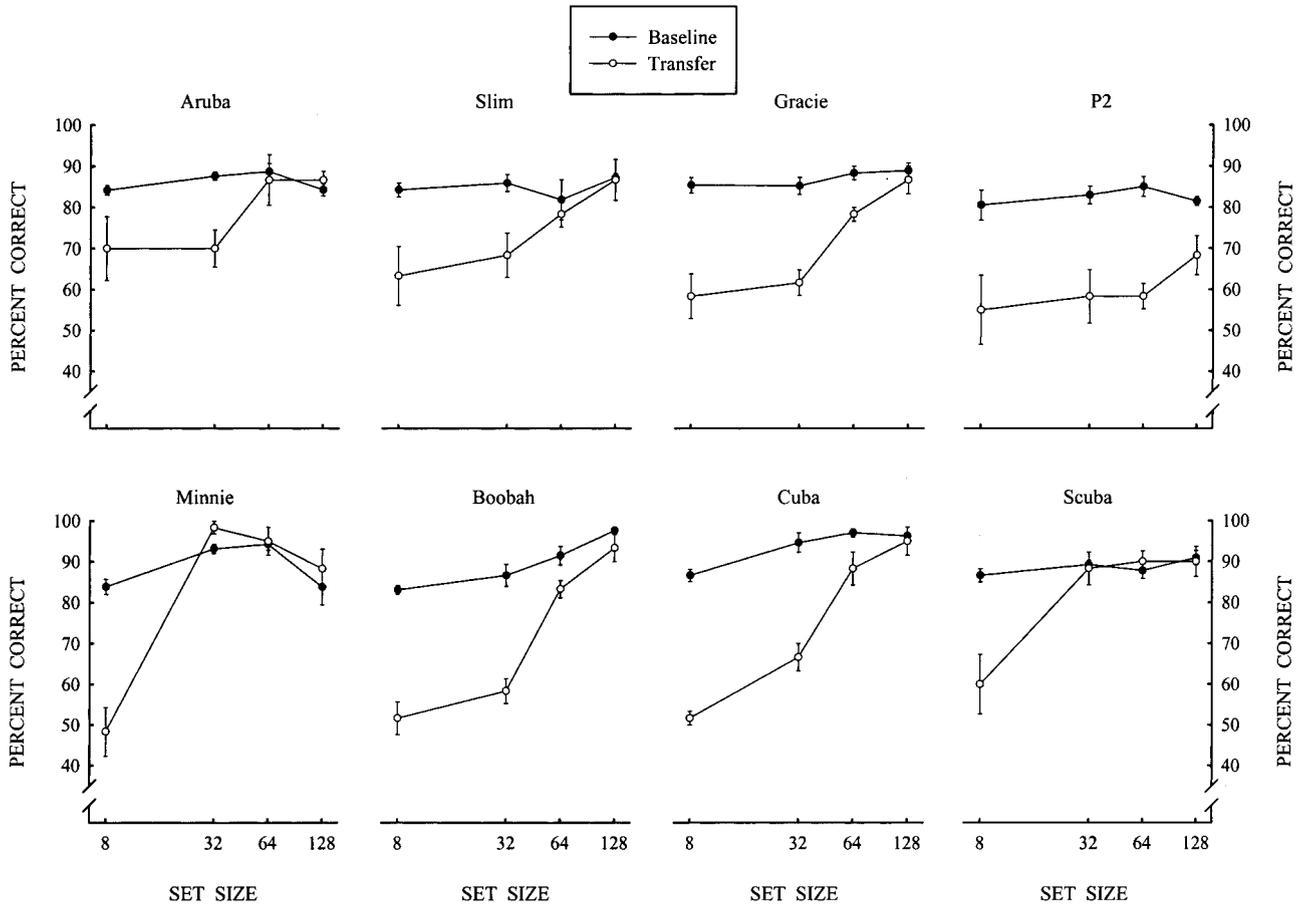
Figure 6. Mean percent correct performance and standard errors for baseline (filled circles) and transfer (unfilled circles) for each monkey in the expanding set-size procedure.

being presented simultaneously (0-response group), or the first item being presented for a fixed time (Experiment 2) equivalent to the average exposure time of the 10-response group. Experiment 3 showed that monkeys in Experiments 1 and 2 that did not learn the

Figure 7. Mean percent correct performance and standard errors for baseline (filled circles) and transfer (unfilled circles) for the control monkey, Cubby, originally trained like the other (experimental) monkeys, Boobah, Cuba, and Scuba, of the 10-response group, but with the set size (8 items) fixed, rather than expanded, following Experiment 1.

task did learn when they were switched to the same 10-response procedure as the 10-response group. Experiment 4 showed that the training stimulus set size was critical to abstract-concept learning. Abstract-concept learning was strongly correlated ($r = .89$) with training set size. After training with the 8-item set, novel-stimulus transfer was 57% correct, very close to chance performance, but transfer progressively rose over four doublings of the set size and ended up equivalent to training baseline performance and 87% correct. Furthermore, full abstract-concept learning was obtained for all monkeys after the training set size was expanded to 128 items except the 1 monkey in the 0-response group.

These manipulations of the observing response requirement and training set size are in stark contrast to experiments designed to test whether some species simply can or cannot learn an abstract concept. For example, early tests with pigeons uniformly failed to find MTS abstract-concept learning (Berryman et al., 1965; Cumming & Berryman, 1961; Cumming et al., 1965; D'Amato et al., 1985; Farthing & Opuda, 1974; Holmes, 1979). Indeed, these failures of MTS abstract-concept learning by pigeons provided a major stimulus for Premack's (1978) article. Among other things, these early experiments testing pigeon MTS abstract-concept learning were handicapped by the small number of training and testing stimuli, which were limited to 12-stimulus projector units in standard use at that time. Consider the conclusion if the exper-

iments of the present article had been stopped after training and testing with the 8-item set. Without evidence to the contrary, it might have been concluded that monkeys cannot learn S/D abstract concepts. Other theories following Premack (1978, 1983) have proposed that abstract-concept learning reflects a species' cognitive capabilities (and its intelligence; D'Amato et al., 1985; Thomas, 1980, 1986, 1996). These cognitive capabilities were thought to be all or none; either they can do it (learn the abstract concept) or they cannot. Thus, some species were thought to be unique (e.g., language-trained chimpanzees) relative to other species in their abstract-concept learning ability. Similar proposals can be found today in claims that chimpanzees are unique in their analogical reasoning ability to solve relations among relations (Thompson & Oden, 2000; but see Bovet & Vauclair, 2001; Fagot, Wasserman, & Young, 2001, for analogical reasoning by baboon monkeys). The present article deviates from this unique-species-ability approach in that we believe the search for qualitative (unique) abstract-concept learning abilities is misdirected, if not futile. We suggest instead that, for the most part, there are quantitative differences in abstract-concept learning and that to identify these quantitative differences an understanding of the processes responsible for abstract-concept learning is needed.

Instead of an all-or-nothing result with regard to abstract-concept learning, we showed in the present article that S/D abstract-concept learning varied with training set size. Thus, the size of the training stimulus set was instrumental in producing S/D abstract-concept learning. There are examples of chimpanzees learning an S/D abstract concept (Robinson, 1955) and an MTS abstract concept (Oden et al., 1988) with small set sizes. But we propose that these differences ultimately may represent quantitative differences, not qualitative differences. The proposal made here is that most species (e.g., most vertebrates) ultimately have a set-size function for abstract-concept learning. Some species (e.g., pigeons) may have a set-size function that is lower than that for rhesus monkeys and would require larger set sizes to achieve full abstract-concept learning. Other species (e.g., chimpanzees) may have a set-size function that is elevated relative to that for rhesus monkeys. If the set-size function is sufficiently elevated, then that species under those conditions might be able to learn an abstract concept with very few items. Humans can demonstrate equivalence relationships after being trained with small stimulus sets (e.g., Sidman et al., 1982). Another example would be chimpanzees that were trained in matching to sample with two objects, a metal lock and metal cup, showed full MTS abstract-concept learning (Oden et al., 1988). Possibly if these chimpanzees had been shown two-dimensional pictures of unfamiliar abstract shapes, artwork, or patterns (e.g., kaleidoscope patterns) presented on a video monitor, they might have not learned so quickly and might have not shown abstract-concept learning with just two training items. Expanding the set size might then have induced abstract-concept learning producing a set-size function similar, but quantitatively different (i.e., elevated), relative to that for rhesus monkeys. Just because some species can learn a particular abstract concept with a small set size does not mean that set size is unimportant for all concept learning for this species. There will likely be some combination of variables and conditions where some abstract concept would not be learned with a small set size, and under such conditions the set size could then be manipulated and its effects on concept learning tested.

So far little has been said about the role of the observing-response requirement in abstract-concept learning. Clearly, contact with the individual items was important for the monkeys learning the S/D task in the present experiment and learning other tasks as well (e.g., Harrison et al., 1977; Meyer et al., 1965; Stollnitz, 1965). Indeed, having real three-dimensional objects to touch and manipulate may have served the function of an observing response requirement in the chimpanzee abstract-concept learning experiment with a 2-item set (Oden et al., 1988). Observing responses, like set size, appears to be an important factor in abstract-concept learning.

In conclusion, the proposal made here emphasizes the similarity in the underlying processes and mechanisms of abstract-concept learning (e.g., set-size function), rather than a qualitative difference among species in their cognitive capabilities (e.g., unique concept learning abilities). The "concept" of unique species abilities asserts that there are qualitative differences among species. By contrast, the proposal made here is that there are quantitative differences among species (e.g., elevated or lowered set-size function). We believe that most species will show similar functional relationships of abstract-concept learning and hence similar mechanisms of concept learning.

References

Berryman, R., Cumming, W. W., Cohen, L. R., & Johnson, D. F. (1965). Acquisition and transfer of simultaneous oddity. *Psychological Reports, 17,* 767–775.

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.

Bhatt, R. S., & Wright, A. A. (1992). Concept learning by monkeys with video picture images and a touch screen. *Journal of the Experimental Analysis of Behavior, 57,* 219–225.

Bovet, D., & Vauclair, J. (2001). Judgment of conceptual identity in monkeys. *Psychonomic Bulletin & Review, 8,* 470–475.

Cook, R. G., Cavoto, K. K., & Cavoto, B. R. (1995). Same–different texture discrimination and concept learning in pigeons. *Journal of Experimental Psychology: Animal Behavior Processes, 21,* 253–260.

Cumming, W. W., & Berryman, R. (1961). Some data on matching behavior in the pigeon. *Journal of the Experimental Analysis of Behavior, 4,* 281–284.

Cumming, W. W., Berryman, R., & Cohen, L. R. (1965). Acquisition and transfer to zero delay matching. *Psychological Reports, 17,* 435–445.

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.

Fagot, J., Wasserman, E. A., & Young, M. E. (2001). Discriminating the relation between relations: The role of entropy in abstract conceptualization by baboons (*Papio papio*) and humans (*Homo sapiens*). *Journal of Experimental Psychology: Animal Behavior Processes, 27,* 316–328.

Farthing, G. W., & Opuda, M. J. (1974). Transfer of matching-to-sample in pigeons. *Journal of the Experimental Analysis of Behavior, 21,* 199–213.

Harrison, J. M., Iversen, S. D., & Pratt, S. R. (1977). Control of responding by location of auditory stimuli: Adjacency of sound and response. *Journal of the Experimental Analysis of Behavior, 28,* 243–251.

Herman, L. M., Hovancik, J. R., Gory, J. D., & Bradshaw, G. L. (1989). Generalization of visual matching by a bottlenosed dolphin (*Tursiops truncatus*): Evidence for invariance of cognitive performance with visual and auditory materials. *Journal of Experimental Psychology: Animal Behavior Processes, 15,* 124–136.

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

Holmes, P. W. (1979). Transfer of matching performance in pigeons. *Journal of the Experimental Analysis of Behavior, 31,* 103–114.

James, W. (1950). *The principles of psychology* (Vol. 1). New York: Dover Publications. (Original work published 1890)

Kastak, D., & Schusterman, R. J. (1994). Transfer of visual identity matching-to-sample in two California sea lions (*Zalophus californians*). *Animal Learning & Behavior, 22,* 427–435.

Meyer, D. R., Treichler, F. R., & Meyer, P. M. (1965). Discrete-trial training techniques and stimulus variables. In A. M. Schrier, H. F. Harlow, & F. Stollnitz (Eds.), *Behavior of nonhuman primates* (Vol. 1, pp. 1–49). New York: Academic Press.

Moon, L. E., & Harlow, H. F. (1955). Analysis of oddity learning by rhesus monkeys. *Journal of Comparative and Physiological Psychology, 48,* 188–195.

Oden, D. L., Thompson, R. K. R., & Premack, D. (1988). Spontaneous transfer of matching by infant chimpanzees (*Pan troglodytes*). *Journal of Experimental Psychology: Animal Behavior Processes, 14,* 140–145.

Overman, W. H., Jr., & Doty, R. W. (1980). Prolonged visual memory in macaques and man. *Neuroscience, 5,* 1825–1831.

Pepperberg, I. M. (1987). Acquisition of the same/different concept by an African Grey parrot (*Psittacus erithacus*): Learning with respect to categories of color, shape, and material. *Animal Learning and Behavior, 15,* 423–432.

Piaget, J., & Inhelder, B. (1969). *The psychology of the child* (H. Weaver, Trans.). New York: Basic Books. (Original work published 1966)

Premack, D. (1978). On the abstractness of human concepts: Why it would be difficult to talk to a pigeon. In S. H. Hulse, H. Fowler, & W. K. Honig (Eds.), *Cognitive processes in animal behavior* (pp. 423–451). Hillsdale, NJ: Erlbaum.

Premack, D. (1983). Animal cognition. *Annual Review of Psychology, 34,* 351–362.

Premack, D., & Premack, A. J. (1983). *The mind of an ape.* New York: Norton.

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

Robinson, J. S. (1955). The sameness–difference discrimination problem in chimpanzee. *Journal of Comparative and Physiological Psychology, 48,* 195–213.

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

Sands, S. F., & Wright, A. A. (1980a, August). Primate memory: Retention of serial list items by a rhesus monkey. *Science, 209,* 938–940.

Sands, S. F., & Wright, A. A. (1980b). Serial probe recognition performance by a rhesus monkey and a human with 10- and 20-item lists. *Journal of Experimental Psychology: Animal Behavior Processes, 6,* 386–396.

Santiago, H. C., & Wright, A. A. (1984). Pigeon memory: Same/different concept learning, serial probe recognition acquisition, and probe delay effects on the serial-position function. *Journal of Experimental Psychology: Animal Behavior Processes, 10,* 498–512.

Sidman, M. S., 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. *Journal of the Experimental Analysis of Behavior, 37,* 23–44.

Stollnitz, F. (1965). Spatial variables, observing responses, and discrimination learning sets. *Psychological Review, 72,* 247–261.

Thomas, R. K. (1980). Evolution of intelligence: An approach to its assessment. *Brain, Behavior and Evolution, 17,* 454–472.

Thomas, R. K. (1986). Vertebrate intelligence: A review of the laboratory research. In R. J. Hoage & L. Goldman (Eds.), *Animal intelligence: Insights into the animal mind* (pp. 37–56). Washington, DC: Smithsonian Institution Press.

Thomas, R. K. (1996). Investigating cognitive abilities in animals: Unrealized potential. *Cognitive Brain Research, 3,* 157–166.

Thompson, R. K. R., & Oden, D. L. (2000). Categorical perception and conceptual judgments by nonhuman primates: The paleological monkey and the analogical ape. *Cognitive Science, 24,* 363–396.

Thompson, R. K. R., Oden, D. L., & Boysen, S. T. (1997). Language-naive chimpanzees (*Pan troglodytes*) judge relations between relations in a conceptual matching-to-sample task. *Journal of Experimental Psychology: Animal Behavior Processes, 23,* 31–43.

Washburn, D. A., Hopkins, W. D., & Rumbaugh, D. M. (1989). Video-task assessment of learning and memory in macaques (*Macaca mulatta*): Effects of stimulus movement on performance. *Journal of Experimental Psychology: Animal Behavior Processes, 15,* 393–400.

Wasserman, E. A., Hugart, J. A., & Kirkpatrick-Steger, K. (1995). Pigeons show same/different conceptualization after training with complex visual stimuli. *Journal of Experimental Psychology: Animal Behavior Processes, 21,* 248–252.

Weinstein, B. (1941). Matching-from-sample by rhesus monkeys and by children. *Journal of Comparative Psychology, 41,* 195–213.

Wright, A. A. (1997). Concept learning and learning strategies. *Psychological Science, 8,* 119–123.

Wright, A. A. (2001). Learning strategies in matching to sample. In R. G. Cook (Ed.), *Avian visual cognition.* Retrieved from http://www.pigeon.psy.tufts.edu/avc/

Wright, A. A., Cook, R. G., Rivera, J. J., Sands, S. F., & Delius, J. D. (1988). Concept learning by pigeons: Matching-to-sample with trial-unique video picture stimuli. *Animal Learning & Behavior, 16,* 436–444.

Wright, A. A., Santiago, H. C., & Sands, S. F. (1984). Monkey memory: Same/different concept learning, serial probe acquisition, and probe delay effects. *Journal of Experimental Psychology: Animal Behavior Processes, 10,* 513–529.

Wright, A. A., Shyan, M. R., & Jitsumori, M. (1990). Auditory same/different concept learning by monkeys. *Animal Learning & Behavior, 18,* 287–294.

Young, M. E., Wasserman, E. A., & Dalrymple, R. M. (1997). Memory-based *same-different* conceptualization by pigeons. *Psychonomic Bulletin & Review, 4,* 552–558.