

SHORT REPORTS

Pigeons Show Same–Different Conceptualization After Training With Complex Visual Stimuli

Edward A. Wasserman, Jacob A. Hugart, and Kim Kirkpatrick-Steger
University of Iowa

Eight pigeons were first trained to peck 1 button in the presence of 16 distinct 4×4 arrays of identical pictures and to peck a 2nd button in the presence of 16 distinct 4×4 arrays of nonidentical pictures. Later, they were tested with 16 other *same* and 16 other *different* stimulus arrays involving untrained pictures. Performance to the testing arrays greatly exceeded chance levels, thus suggesting same–different conceptualization by pigeons.

A nickel, a dime, and a quarter. Three dimes. A knife, a fork, and a spoon. Three forks. When the first author recently asked his 3-year-old son David which of these four novel arrays of objects contained items that were all the same as one another or all different from one another, David correctly responded, “different,” “same,” “different,” and “same.” Most behavior theorists such as Herrnstein (1990) interpret performance like David’s as a representation of abstract conceptual behavior. After training with limited aggregations of items, the responses *same* and *different* can be applied to limitless aggregations of items that conform to certain interitem relations. Of course, the above vignette is only one of many thousands that testify to the ability of human beings—even very young ones—to exhibit abstract conceptual behavior. The case for nonhuman animals is not nearly so clear.

The search for abstract conceptualization by nonhuman animals has been a long and arduous one, and it has spanned a century of research in animal intelligence (Wasserman, 1993). Researchers have explored different behavioral methods and animal species in this search, with mixed results. With some combinations of tasks and species, empirical evidence of abstract conceptual behavior has been obtained; however, with other combinations of tasks and species, empirical evidence of abstract conceptual behavior has not been obtained.

There is a growing appreciation that most laboratory investigators of conceptual behavior in nonhuman animals have probably underestimated the cognitive capacities of their subjects (see Wright, 1992, and Zentall, 1993, for further discussion and analysis of this issue). Some authors

have questioned the overly restrictive use of a very few experimental tasks like matching-to-sample (e.g., Dube, McIlvane, & Green, 1992; Lombardi, Fachinelli, & Delius, 1984; Macphail & Reilly, 1989). Other authors have questioned the use of very small numbers of extremely simple stimuli in these few tasks (e.g., Santiago & Wright, 1984). Either or both of these limiting factors might lead workers to arrive at an inappropriately dim view of animal cognition.

Our project tried to take to heart these well-justified concerns with most conventional studies of abstract conceptualization by animals. We used relatively large numbers of rather complex visual stimuli and a little exploited, but very promising, same–different discrimination task (cf. Edwards, Jagielo, & Zentall, 1983; Santiago & Wright, 1984) in our effort to document abstract conceptual behavior in pigeons. Our particular implementation of this experimental task took advantage of current computer and touch-screen technology, and it closely conformed with the very kind of same–different conceptualization by human beings described at the beginning of this article, in which the visual discriminative stimuli involved multiple copies of the same item or single copies of different items.

We trained our 8 pigeons with 32 distinct visual stimuli, each of which comprised 16 small drawings (so-called icons from the Macintosh computer system) in a 4×4 array. Half of the composite stimuli (those that we termed *same*) comprised 16 copies of the identical icon on the pigeon’s video screen. Because there were 16 individual icons in the total pool for constructing all of the training stimuli, there were 16 distinct *same* arrays. The other half of the composite stimuli (those that we termed *different*) comprised each of the 16 icons; the 16 distinct *different* arrays contained each of the icons in various venues, so that across all 16 composite stimuli, each icon was located in each of the 16 regions of the 4×4 array. (The upper half of Figure 1 illustrates a sampling of the *same* and *different* arrays that a pigeon might have seen during its training sessions.) The pigeons’ task during training was simply to peck one button in the presence of each of the 16 *same* arrays and to peck a second button in the presence of each of the 16 *different* arrays to receive food reinforcement.

Edward A. Wasserman, Jacob A. Hugart, and Kim Kirkpatrick-Steger, Department of Psychology, University of Iowa.

This research was supported by National Institute of Mental Health Grants MH 47313 and 51562.

Correspondence concerning this article should be addressed to Edward A. Wasserman, Department of Psychology, University of Iowa, Iowa City, Iowa 52242-1407. Electronic mail may be sent via Internet to ed-wasserman@uiowa.edu.

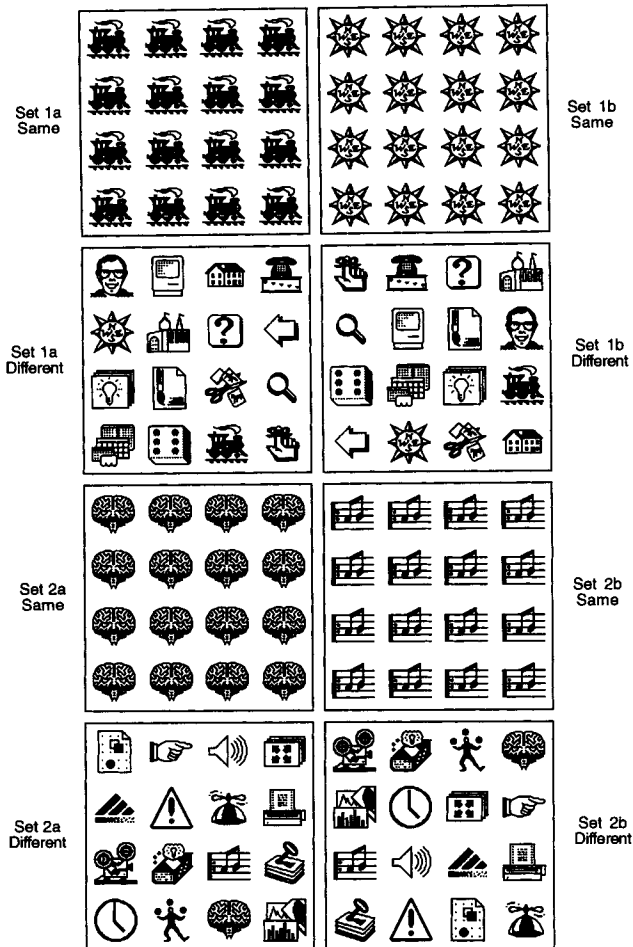


Figure 1. A sampling of eight of the sixty-four 16-icon arrays used in the study. These 64 arrays were assigned to four subsets of 16 icons each (8 *same* and 8 *different* arrays): Set 1a, Set 1b, Set 2a, and Set 2b. See the text for a complete explanation of the creation and function of these subsets of *same* and *different* stimulus arrays.

As the critical test for conceptualization, after discrimination training produced high levels of choice accuracy, we showed the pigeons another set of 32 composite stimuli—16 of the *same* variety and 16 of the *different* variety—in whose presence the birds had never before been given differential food reinforcement. (The lower half of Figure 1 illustrates a sampling of the *same* and *different* arrays that a pigeon might have seen during its testing sessions.) Accurate choice responding in the presence of these 32 testing arrays constituted the prime evidence of same–different conceptualization.

We did in fact observe highly discriminative responding to the testing arrays, leading us to conclude that pigeons can come to respond in accord with the abstract concept that collections of numerous complex visual stimuli are all the same or that they are all different.

Method

Subjects

The subjects were 8 feral pigeons maintained at 85% of their free-feeding weights by controlled daily feeding. The pigeons had served earlier in unrelated studies involving different visual stimuli in standard Skinner boxes that were outfitted with microswitch-activated pecking keys and solenoid-operated grain hoppers. Before this experiment, the birds were taught to peck at the touch screen and to consume the food pellets delivered by the rotary dispenser.

Apparatus

The experiment used four custom-built chambers. One plywood wall of each chamber contained a large opening with an aluminum frame attached to the outside of the box. The frame held a clear touch screen (Elmwood Sensors DuraTouch Model #70056-001; see Allan & Zeigler, 1989, for more on the use of touch technology in animal behavior research); pecks on the touch screen were processed by a serial controller board (Elographics Model #E271-2200). A brushed aluminum panel was placed directly in front of the touch screen to allow the pigeons access to circumscribed portions of a video monitor (13-in. AppleColor™ High-Resolution RGB) that was located 0.9 cm behind the touch screen from the monitor's center and 1.1 cm from the outer edges of the monitor (the difference being due to the slight convex curvature of the face of the monitor). The panel emulated one used by Bhatt, Wasserman, Reynolds, and Knauss (1988) in their studies of basic concept learning by pigeons. There were five openings or buttons in the panel: a 7-cm × 7-cm central "display" area in which the icon arrays appeared and four round "report" areas (1.9-cm diameter) located 2.3 cm from each of the four corners of the central opening. Only the lower two report areas were used; the lower left area could be lit red and the lower right area could be lit green. A clear Plexiglas food cup was centered on the rear wall of the chamber to discourage the pigeons from perching on the food cup; it was recessed into the wire mesh floor so that the top of the cup was level with the floor. A pellet dispenser (MED Associates Model #ENV-203M) delivered 45-mg Noyes pigeon pellets through a vinyl tube into the food cup. A houselight, mounted on the upper rear wall of the chamber, provided constant illumination during experimental sessions. The houselight and pellet dispenser were controlled by a digital input/output (I/O) interface board (National Instruments Model #NB-DIO-24).

In two of the chambers, the control of the peripheral stimuli (through the I/O interface card) and the recording of the subjects' responses (through the serial controller board) were accomplished by two Apple Macintosh IIci computers. A video splitter (Network Technologies Model #Vopex 2M) connected each computer to the pigeon's monitor and to an identical monitor located in an adjacent room. In the other two chambers, the control of the peripheral stimuli and the recording of the subjects' responses were accomplished by two Apple Macintosh Quadra 650 computers. A distribution amplifier (Extron Model #MAC/2 DA2) connected each computer to the pigeon's monitor and to an identical monitor located in an adjacent room. Programs were developed in HyperCard (Version 2.2).

Visual Stimuli and Experimental Design

Thirty two highly distinguishable Macintosh icons were chosen as the total item pool; these icons were randomly sorted into two

sets of 16 icons each (Set 1 and Set 2), and from these two 16-icon sets, sixty-four 16-item arrays were constructed (32 *same* and 32 *different*, with 16 *same* and 16 *different* arrays composed from icons in each of Sets 1 and 2). The main experimental plan was first to train 4 birds with the 16 *same* and 16 *different* arrays of icons from Set 1 and then to test those birds with the 16 *same* and 16 *different* arrays of icons from Set 2. The other 4 birds were first to be trained with the 16 *same* and 16 *different* arrays of icons from Set 2 and then to be tested with the 16 *same* and 16 *different* arrays of icons from Set 1.

Before implementing that plan—which was actually the second round of training and testing—we first trained the 4 birds in each group of subjects with only half of the stimulus arrays in the full training set. For example, 2 birds in the first group of 4 pigeons were initially trained with the eight *same* and eight *different* stimulus arrays that were randomly assigned to Set 1a; then they were tested with the remaining eight *same* and *different* stimulus arrays in Set 1b and with the 16 *same* and 16 *different* stimulus arrays in Set 2. We reasoned that testing arrays that entailed icons in the training set (e.g., Set 1b) might, because of their greater familiarity, support higher levels of transfer performance than would testing arrays that did not involve any of the icons in the training set (e.g., Set 2). Figure 1 illustrates a sampling of the stimulus arrays in Set 1a, Set 1b, Set 2a, and Set 2b. These two preliminary stages constituted the first round of training and testing (whose results will be seen not to agree with our initial expectations).

Procedure

Training 1. These training sessions comprised 160 daily trials: 10 randomized blocks of 16 trials each, 8 with *same* arrays and 8 with *different* arrays. For four pairs of birds, the 16 training arrays came from Set 1a, Set 1b, Set 2a, and Set 2b, respectively. Pecks to the red key on *same* trials and to the green key on *different* trials were correct and were reinforced with one or two pellets of food (depending on each bird's 85% weight). Pecks to the green key on *same* trials and to the red key on *different* trials were incorrect and were punished by repetition of the trial until the correct response was made.

Each trial began with illumination of the display area with a white field containing a black boxed X in the center. A single peck anywhere in the display area turned on the 4 × 4 icon array as a black-on-white picture. After a fixed number of pecks (*observing responses*) anywhere in the display area (the number of pecks ranged from 10 to 35 for individual pigeons, depending on the bird's tolerance of high ratios), the icon array reversed to a white-on-black picture (to alert the bird to the availability of the report buttons), and the lower two report buttons were lit red (left) and green (right). A correct-choice response blackened the display area and the report buttons and delivered food pellet reinforcement; an incorrect-choice response blackened the report buttons, returned the picture to its black-on-white state, and began a series of one or more correction trials that were not scored in data analyses. Intertrial intervals averaged 10 s (range = 5–15 s). Training 1 continued for 50 sessions, by which time the pigeons were averaging about 85% correct responses.

Testing 1. Four testing sessions were given, each of which was separated by two sessions of retraining that were procedurally identical to Training 1. Testing sessions began with 32 warm-up trials with each of the 16 training arrays shown twice; thereafter, 2 randomly selected training trials were given followed by 1 randomly selected testing trial until the session ended and each of

the testing stimuli was seen once. Over the last 144 trials of the testing sessions, 96 training trials were given, 16 testing trials with novel arrays of icons from the training set were given, and 32 testing trials with novel arrays of icons not from the training set were given. Differential food reinforcement for left and right report responses was given on training trials. Nondifferential reinforcement for left and right report responses was given on testing trials; food was given regardless of the pigeons' choice responses, so that repeated testing could be conducted without confounding by explicit training effects (Wasserman, DeVolder, & Coppage, 1992). Thus, the designations *correct* and *incorrect* for choice responses on testing trials were for scoring purposes only. No correction trials were programmed on testing trials.

Training 2. These sessions were just like those in Training 1 except that all of the 32 icon arrays in the training set were given with differential food reinforcement. Daily sessions in Training 2 involved 160 trials: five blocks of 32 trials with all of the 16 *same* and 16 *different* arrays in the training set. Training 2 continued for 28 sessions, by which time the pigeons were averaging about 83% correct responses.

Testing 2. These four successive daily testing sessions were similar to those in Testing 1. Testing sessions began with 64 warm-up trials, with each of the 32 training arrays shown twice; thereafter, 2 randomly selected training trials were given followed by 1 randomly selected testing trial until each of the testing arrays had been seen once. Over the last 96 trials of the testing sessions, 64 training trials were given plus 32 testing trials involving those arrays of icons that had not come from the training set and that had been shown just four times each 28 days earlier in Testing 1. Differential food reinforcement for left- and right-button responses was given on training trials; nondifferential food reinforcement for left- and right-button responses was given on testing trials.

Results

Testing 1

Across all four sessions of Testing 1, discriminative performance to the training arrays was very high, averaging 87% correct overall (88% correct for *same* arrays and 85% correct for *different* arrays) over the last 144 trials of the testing sessions. Overall accuracy to the novel arrays composed of familiar (training set) icons averaged 52% correct. Accuracy to these arrays did not differ reliably from the chance score of 50% on any of the four sessions of testing (two-tailed binomial $p > .05$). Overall accuracy to the novel arrays composed of unfamiliar (nontraining set) icons averaged 64% correct. Accuracy to these arrays did differ reliably from chance on each of the four sessions of testing (two-tailed binomial $p < .01$). The overall accuracy scores are depicted in Figure 2.

Although the overall accuracy scores to the novel arrays composed of unfamiliar icons did differ reliably from chance, the accuracy scores to the *same* and *different* stimulus arrays were quite dissimilar: Accuracy to the *same* arrays averaged 46% correct, whereas accuracy to the *different* arrays averaged 82% correct. This dissimilarity was further reflected by the fact that the accuracy scores on each of the four testing sessions were significantly different from chance for the *different* arrays (two-tailed binomial $p < .01$) but not for the *same* arrays (two-tailed binomial $p > .05$).

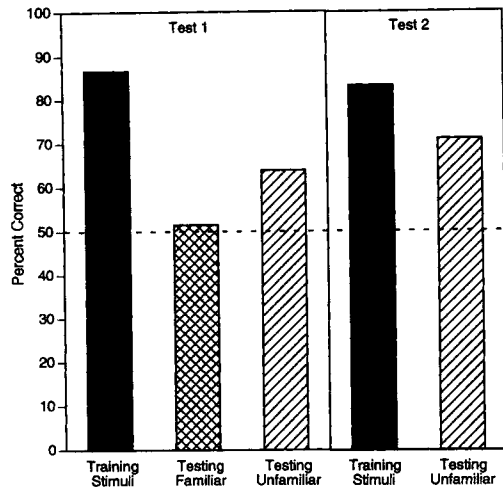


Figure 2. The mean percentage of correct-choice responses to the training arrays (Training Stimuli), to the testing arrays that were composed from the icons shown in the training arrays (Testing Familiar), and to the testing arrays that were composed from the icons not shown in the training arrays (Testing Unfamiliar) during the two different phases of testing (Test 1 and Test 2).

An even wider performance differential separated the accuracy scores on *same* and *different* trials for the novel arrays composed of familiar icons: Accuracy to the *same* arrays averaged 28% correct, whereas accuracy to the *different* arrays averaged 75% correct. This differential meant that the below-chance accuracy scores on each of the four testing sessions were significantly different from chance for the *same* arrays (two-tailed binomial $p < .05$) and that above-chance accuracy scores on each of the four testing sessions were significantly different from chance for the *different* arrays (two-tailed binomial $p < .01$).

One way to interpret these testing results is to hypothesize that initial discrimination training may have disposed the pigeons to respond *same* to only the 8 *same* arrays with which they were trained and to respond *different* to any other stimulus arrays. This disposition would have produced a preponderance of right-button responses on the new *same* and *different* trials during testing, because all of the testing arrays differed from the 8 *same* arrays in the training set, thereby producing below-chance scores on *same* trials (in which right report responses would have been scored *incorrect*) and above-chance scores on *different* trials (in which right report responses would have been scored *correct*). Why this disposition should have been stronger on trials with testing arrays composed from the training icons than on trials with testing arrays not composed from the training icons is unclear. However, Testing 2 involved no testing arrays that were composed from the training icons. It also followed Training 2, which entailed 16 *same* arrays and 16 *different* arrays—twice as many as were given in Training 1. Memorization of the specific arrays in the *same* set should thus have been less likely than in Training 1.

Testing 2

Across all four sessions of Testing 2, discriminative performance on the training arrays was very high, averaging 83% correct overall (86% to *same* arrays and 81% to *different* arrays) over the last 144 trials of the testing sessions. Overall accuracy to the testing arrays composed only from untrained icons averaged 71% correct, which differed reliably from chance on each of the four sessions of testing (two-tailed binomial $p < .01$). The overall accuracy scores are depicted in Figure 2. The accuracy scores to the *same* and *different* testing arrays were now quite similar: Accuracy to the *same* arrays averaged 70% correct, and accuracy to the *different* arrays averaged 72% correct. This similarity was further reflected in the fact that the accuracy scores on each of the four testing sessions were significantly different from chance for both the *same* and *different* testing arrays (two-tailed binomial $p < .01$).

Discussion

Our results show that the pigeons could generalize the *same* and *different* report responses that they learned with one set of 16 distinct arrays of identical icons and 16 distinct arrays of nonidentical icons to another set of *same* and *different* arrays comprising icons in whose presence the birds had never before been given differential food reinforcement. The obtained mean of 71% correct responses to these testing arrays (in Testing 2) after training with 16 *same* and 16 *different* arrays (in Training 2) was not only substantially above the chance level of 50% correct, but it exceeded the obtained mean of 64% correct responses to those very testing arrays (in Testing 1) after training with a set of only 8 *same* and 8 *different* arrays (in Training 1). This finding closely corresponds with other research on natural and abstract conceptualization by humans and non-humans showing that larger sets of training stimuli engender stronger generalization performance in testing (see Wasserman, 1993, for a review of these data).

Whether training with even larger sets of *same* and *different* arrays of icons will support even higher levels of generalization performance to untrained stimulus arrays is a question to be further researched. So too is the question of whether our pigeons' stronger generalization performance after learning with a set of 32 arrays (in Training 2) depended on their prior learning with a set of only 16 arrays (in Training 1). The effectiveness of our methodology clearly encourages us to pursue these and other interesting questions in the pigeon's acquisition of abstract conceptual behavior.

For example, is discrimination learning and generalization testing performance affected by varying the number of items in the stimulus arrays or by permitting or prohibiting the training icons from appearing in both the *same* and *different* stimulus arrays? Seeking answers to these latter questions would help us to better relate our research to the earlier work of Santiago and Wright (1984), who first trained their pigeons with only 2 visual items (color photographs) in each array and who did not permit the elements

of these 2-item arrays to appear on both *same* and *different* trials. We trained our pigeons with 16 visual items (Macintosh icons) in each array and did permit the elements of these arrays to appear on both *same* and *different* trials.

This search might also help us to better understand the role of perceptual processes in our birds' conceptual behavior. With arrays of 16 items, "sameness" and "differentness" would appear to be far more obvious than with arrays of only 2 items. Indeed, it was this very obviousness that led us to suspect that our task stimuli might more strongly support abstract conceptualization than prior task stimuli (cf. Edwards et al., 1983; Santiago & Wright, 1984). However, could our visual stimuli have actually engaged some low-level perceptual process that might explain our pigeons' seemingly advanced discrimination and generalization behavior?

The role of perceptual processes in conceptual behavior is not to be underestimated (Cook, Cavoto, & Cavoto, 1995; Wasserman & Astley, 1994). Yet, it is difficult to imagine that the pigeons could tell if our visual displays involved identical icons without their also noting the nature of those icons. In fact, a low-level texture detection mechanism cannot by itself explain why the unfamiliar displays in testing led to much less accurate discriminative responding than did the training displays (Figure 1).¹ This generalization decrement suggests that the birds did discriminate specific icons in the visual arrays as well as their relation to one another.

These considerations notwithstanding, it is becoming increasingly clear that the pigeon is capable of evidencing advanced forms of conceptual behavior. Conceptual behavior appears to involve relatively concrete basic-level (Bhatt et al., 1988) and higher-order (Wasserman et al., 1992) categories as well as much more abstract stimulus relations (Cook et al., 1995; Edwards et al., 1983; Santiago & Wright, 1984; the present results), thus arguing against Premack's (1983) earlier conclusion that, among nonhuman animals, only language-trained chimpanzees can show same-different conceptualization.

¹ Analysis of variance on the discrimination scores in Testing 2 (in which generalization performance to the unfamiliar testing arrays was strongest) yielded a significant main effect of training versus testing arrays, $F(1, 7) = 20.59, p < .01$, but no significant main effect of *same* versus *different* display types, $F(1, 7) < 1$, and no significant interaction between these variables, $F(1, 7) < 1$.

References

Allan, R. W., & Zeigler, H. P. (1989). Measurement and control of pecking response location in the pigeon. *Physiology and Behavior*, *45*, 1215-1221.

- Bhatt, R. S., Wasserman, E. A., Reynolds, W. F., Jr., & 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.
- Cook, R. G., Cavoto, K. K., & Cavoto, B. R. (1995). Same-different texture discrimination and concept learning by pigeons. *Journal of Experimental Psychology: Animal Behavior Processes*, *21*, 253-260.
- Dube, W. V., McIlvane, N. J., & Green, G. (1992). An analysis of generalized identity matching-to-sample test procedures. *Psychological Record*, *42*, 17-28.
- Edwards, C. A., Jagielo, J. A., & Zentall, T. R. (1983). "Same/different" symbol use by pigeons. *Animal Learning & Behavior*, *11*, 349-355.
- Herrnstein, R. J. (1990). Levels of stimulus control: A functional approach. *Cognition*, *37*, 133-166.
- Lombardi, C. M., Fachinelli, C. C., & Delius, J. D. (1984). Oddity of visual patterns conceptualized by pigeons. *Animal Learning & Behavior*, *12*, 2-6.
- Macphail, E. M., & Reilly, S. (1989). Rapid acquisition of a novelty versus familiarity concept by pigeons (*Columba livia*). *Journal of Experimental Psychology: Animal Behavior Processes*, *15*, 242-252.
- Premack, D. (1983). The codes of beast and man. *Behavioral and Brain Sciences*, *6*, 125-167.
- 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.
- Wasserman, E. A. (1993). Comparative cognition: Beginning the second century of the study of animal intelligence. *Psychological Bulletin*, *113*, 211-228.
- Wasserman, E. A., & Astley, S. L. (1994). A behavioral analysis of concepts: Its application to pigeons and children. In D. L. Medin (Ed.), *Psychology of learning and motivation* (pp. 73-132). San Diego, CA: Academic Press.
- Wasserman, E. A., DeVolder, C. L., & Coppage, D. J. (1992). Nonsimilarity-based conceptualization in pigeons via secondary or mediated generalization. *Psychological Science*, *3*, 374-379.
- Wright, A. A. (1992). Testing the cognitive capacities of animals. In I. Gormezano & E. A. Wasserman (Eds.), *Learning and memory: The behavioral and biological substrates* (pp. 45-60). Hillsdale, NJ: Erlbaum.
- Zentall, T. R. (1993). Animal cognition: An approach to the study of animal behavior. In T. R. Zentall (Ed.), *Animal cognition: A tribute to Donald A. Riley* (pp. 3-15). Hillsdale, NJ: Erlbaum.

Received September 8, 1994

Revision received December 8, 1994

Accepted December 13, 1994 ■