

Levels of stimulus control: A functional approach*

R.J. HERRNSTEIN
Harvard University

Abstract

Herrnstein, R.J., 1990. Levels of stimulus control: A functional approach. *Cognition* 37: 133–166.

This paper surveys some illustrative experiments on categorization of visual stimuli by animals other than human. The results suggest a classification of categorical powers in five steps from simple discrimination to rote and open-ended categorization, to concepts and the use of abstract relations. Nonhuman animals evidently readily categorize up to the fourth level as here defined, which is the level of concepts. With difficulty, they can sometimes be induced to rise even to the level of abstract relations. It is at the level of abstract relations that a large gap opens up between human categorizations and categorization by other animals.

Introduction

There have been numerous efforts to systematize the stimulus control of behavior, or, as it is more often called outside the behaviorist tradition, categorization.¹ In the spirit of cognitive psychology, for example, are the contributions of Smith and Medin (1981) or the essay by Homa (1984). Pre-

*For support of the preparation of this article and some of the work described herein, Grant IST-85-11606 from the National Science Foundation to Harvard University is gratefully acknowledged. Versions were presented at the Neurosciences Institute Symposium on Signal and Sense, June, 1987; in October, 1987, at the Columbia University Seminar on the Psychobiology of Animal Cognition and Harvard University's "Pigeon Meeting," in April, 1988, to a seminar at the Center for Biological Information Processing at MIT, and in March, 1989, to a seminar at the Department of Psychology at New York University. A largely overlapping version of this article is in G.M. Edelman, W.E. Gall, and W.M. Cowan (Eds.) (in press), *Signal and sense: Local and global order in perceptual maps*, New York: Wiley, under the title "Levels of categorization." Thanks to sponsors and participants of these seminars for eliciting, then shaping, these ideas. Thanks also to M.L. Commons, C.R. Gallistel, R.D. Luce, M. Rilling, H.S. Terrace, W. Vaughan, Jr., G. Zuriff, and an anonymous reviewer, for many helpful comments and criticisms. Reprint requests should be addressed to: Dr. Richard Herrnstein, Psychology Department, Harvard University, 33 Kirkland Street, Cambridge, MA 02138, U.S.A.

¹The review of the literature contained in this paper was intended to be illustrative, not exhaustive. It was largely completed in mid-1987, and, except for a few studies that seem to be especially relevant, I have not attempted to update it for more recent contributions.

mack (1986) proposes a more behavioristic framework, as adapted by him in light of modern psycholinguistic theory. Dore and Dumas (1987) apply the Piagetian theory, with its stages of cognitive development, to subhuman categorical behavior. Some approaches are informed by inferences about what nervous systems do, inferences that may shift radically depending on the level of neurological analysis chosen or on the growth and development of neuroscience – others, by the results of attempting to simulate natural categorization on computers, or to express it in formal theory. Still other approaches attempt to follow developmental pathways (e.g., Commons, Hal-
linan, Fong, & McCarthy, 1990). What follows below is an attempt of my own. Neither physiology, computer simulation, nor specific psychological theory, beyond the general notions of reinforcement theory, explicitly influence the scheme I propose. If the scheme has any virtue at all, it is that it seems to provide a convenient framework for thinking about behavioral data.

The present classification of stimulus control is *functional*, in the sense suggested by a consideration of an animal's stimulus problem in nature. In order to behave appropriately an animal needs a description of the stimuli in any environment that renders as invariant as feasible the predictors of psychologically significant consequences of behavior, which may be called the contingencies of reinforcement (for inherited responses no less than for learned ones – see Herrnstein, 1977). Different occasions may thus dictate different descriptions and different dimensions of invariance. The evidence suggests that natural selection has equipped animals with considerable adaptations for dealing with the categorization problem in this very sense, as will be illustrated below.

The problem gets tricky mainly for two reasons – the “occasion” an animal confronts may, first, bear a logically complex relationship to the physical events of which it is composed and, second, it is in the nature of the relevant physical events to be variable and nonrecurrent. Logical complexity arises because the consequences of behavior rarely depend on the mere presence or absence of features of objects – their precise shapes or sizes or colors or the like – but on more complex concatenations of stimulus descriptors. Similarly, even the ordinary objects in an animal's life are variable and ambiguous. Is the present stimulus an acorn, a stone, a predator's spoor? None of those ordinary objects, nor any other, presents itself to a perceiver invariantly. The incentive to identify as such acorns, stones, and so on, arises from the psychological consequences associated with them – one may eat acorns, flee predators, and ignore stones.

If the psychological consequences have been potent and stable long enough in a species' history, then, by virtue of natural selection, the relevant sensory system may have evolved so that the recognition may have innate elements

in it. If not, recognition is likely to depend mostly on learning. There must be a continuum of innateness, along which an animal's identifications spread out. How they do so is a rich issue for analysis, but, for the question of categorization itself, nothing inherent seems to depend upon whether the association between stimulus and consequence is innate or must be learned. In either case, animals solve categorization problems by analyzing stimuli in ways that remain unaccounted for neurophysiologically and out of the reach of formal theory, let alone computer simulation.

For one species, acorns may themselves segregate into multiple classes; for another, they may be classed with beech nuts. Nothing about the appearance of acorns or beech nuts per se determines how they will be grouped. Categorization problems are jointly determined by physical variation in stimuli and the consequences of behavior. The recognition problem is complex or simple from the animal's standpoint depending on the relations among the stimulating objects, the consequences of action in their presence, and the perceptual machinery of the creature. That is to say, if acorns vary widely given an animal's perceptual apparatus, then the animal faces a more difficult recognition problem than if they vary little (assuming at this point that acorns set a common contingency for its behavior, no matter what they look like). Or, the contingencies may be conditional, depending not on any particular object, but, for example, on conjunctions and disjunctions of particular configurations (spatial or temporal) of objects.

The interaction between objective contingencies of reinforcement and the capacities of animals suggests five levels of categorization, roughly increasing in abstractness from 1 to 5 below. Where in this progression a given animal is operating for a given categorization problem – the operational tests for these levels of categorization – will be discussed at greater length subsequently.

(1) *Discrimination*: The allowable variation in stimuli, given the contingency of reinforcement, may be so small that categorization collapses on to mere discrimination. An animal in a psychological experiment may be reinforced for responding to a patch of light of a particular wavelength, plus some small band of error. Categorization is then merely failure to discriminate. The situation may be such that if the subject can discriminate among particular stimuli, then those stimuli have discriminable behavioral consequences; if the subject cannot discriminate, then they have essentially the same consequences. In nature, this may be a null class, but, in the contrived environments of the psychological laboratory, it has been the standard for studying stimulus control. Although there are worthy questions of discriminative fineness or coarseness involved here, little will be said about this most

elementary level of categorization, beyond noting that it sets the minimum in abstractness, as that term will be used here.

(2) *Categorization by rote*: The stimuli within a class that is implicitly defined by a contingency of reinforcement may vary substantially more than the resolving power of discrimination, but the entire class can be approximated by a relatively small number of discriminable exemplars. By a relatively small number of exemplars, I mean a set of exemplars that a subject could conceivably learn as a list (disregarding any implication of serial organization). The jacks in a deck of cards would be an example of this sort of category, except insofar as jacks resemble each other. Because they may resemble each other, the subject, when recognizing the four jacks in a given deck, may actually be drawing on a category with unlimited membership, namely the category of all jacks in all decks, whether ever seen before or not. The stops on the subway system's "Red Line" in and around Boston are a clearer example. Here, the 23 exemplars can be learned only as a list, assuming there is no underlying rule that determines membership. Given the name "Kendall," or a photograph of the station bearing this name, one either does or does not know that this is a stop on the line. This would be the pure rote level of categorization.

(3) *Open-ended categories*: The distinction just made between jacks and stations on the Red Line suggests the next step in abstractness. If the contingency of reinforcement entails open-ended or very large membership, as jacks do, then some principle beyond pure rote learning is required. Given the variation of objects in nature, and the variation in organisms' vantage points in perceiving them, the typical natural situation brings us to at least this level. Acorns vary; even a given acorn varies in the aspect it presents to a given organism. Organisms at any level of psychological development may be expected to have means at least for sorting exemplars as open-ended categories.

Perceptual similarity appears to be an evolutionary adaptation to contingencies of reinforcement that circumscribe classes with unlimited membership. No account of categorization by animals can fail to recognize the existence of some primitive ordering of perceptual distances between objects. Acorns vary, but they probably resemble each other from the standpoint of any creature that has truck with them as a class. What is, in a purely physical sense, a class with unlimited membership may become, because of similarity, a class containing one or several clusters of mutually similar members. Although these clusters must have their physical descriptions, it is often not the case that psychological proximity is captured by physical proximity along the standard physical variables.

For a creature that deals with acorns and beech nuts as a single class, it is not necessary to argue that the two sets of objects coalesce perceptually. Rather, for this hypothetical creature, the acorn-beech-nut class may comprise two subcategories, two “clusters,” as that term was just used. The creature may be capable of detecting differences not only between acorns and beech nuts, but among examples of either. Similarity is not equivalent to indiscriminability. Objects said to be similar are usually discriminable. Similarity is best viewed as biological, rather than physical: an evolutionary adaptation to the open-endedness of exemplars in the reinforcement contingencies that creatures usually must contend with.

(4) *Concepts*: Level 3, by introducing similarity, converts itself to something resembling Level 2, which is pure rote. By allowing the psychological representation of exemplars some range of variation and a psychological metric of similarity, we rise from rote to open-ended categorization. Nothing has yet been said about any psychological representation of the contingency of reinforcement with respect to which the exemplars are behaviorally invariant. Any such representation would be a concept, as that term will be used here.² A creature that has concepts in this sense would adapt to changing contingencies more efficiently than one that did not. For example, suppose acorns became bitter. It can be argued that an animal with some sort of psychological representation for the *class* of acorns, beyond the exemplars themselves, would associate bitterness with acorns in general, not just with the previously tasted acorns or acorns that look like the previously tasted ones, more rapidly and more completely than one without it. This rapid propagation over class members signifies the concept level of categorization. As will be shown later, it is often hard to test whether an animal is sorting by a concept or by just an open-ended category based on similarity, but, in principle, concepts need involve no similarities at all.

(5) *Abstract relations*: In much of the current work on categorization, the focus is on trying to resolve the exemplars of various categories into features or attributes or values along various dimensions, or the like. It is, after all, self-evident that an acorn is not a unitary perceptual entity (for humans), but that it is somehow composed of finer elements – shape, color, size, and so forth. The elements may be complex and relational, including disjunctions, conjunctions, and conditional properties of all sorts. However, this last level

²Few terms have been used in psychological and epistemological theory in as many different senses as “concept” and its derivatives. I will make no attempt to review those senses, nor to relate them to my proposed usage. Some of this is done in Gallistel’s accompanying introduction.

of categorization deals not with the exemplars themselves, but with relations between and among concepts, as defined at Level 4. A reinforcement contingency stipulating “acorns on stumps” exemplifies an abstract relation as meant here, assuming that acorns and stumps are represented as concepts. Another example would be: “color chips in stacks of the same colors.” The relations are carried by “on” or “the same,” respectively. The abstractness refers to the idea that it is not particular acorns or stumps or colors that are characterized, but acorns, stumps, and colors in general, which is to say, the concepts that embrace them. Creatures capable of forming concepts may be incapable of responding differentially and reliably to such relations between concepts when they are tested with new exemplars. But if they can so respond, they would be categorizing according to abstract relations.

Interest in categorization is often motivated by an interest in language, or human cognition more generally. For humans, language is the natural behavioral medium for displaying categories, particularly those beyond the lowest levels of abstraction defined above. The foregoing list stops about where psycholinguistic considerations become interesting, namely, where language embodies not just concepts and relations between concepts, but relations between relations, concepts whose exemplars are themselves abstract relations, intentions and other descriptions of a speaker’s psychological state, and so on.

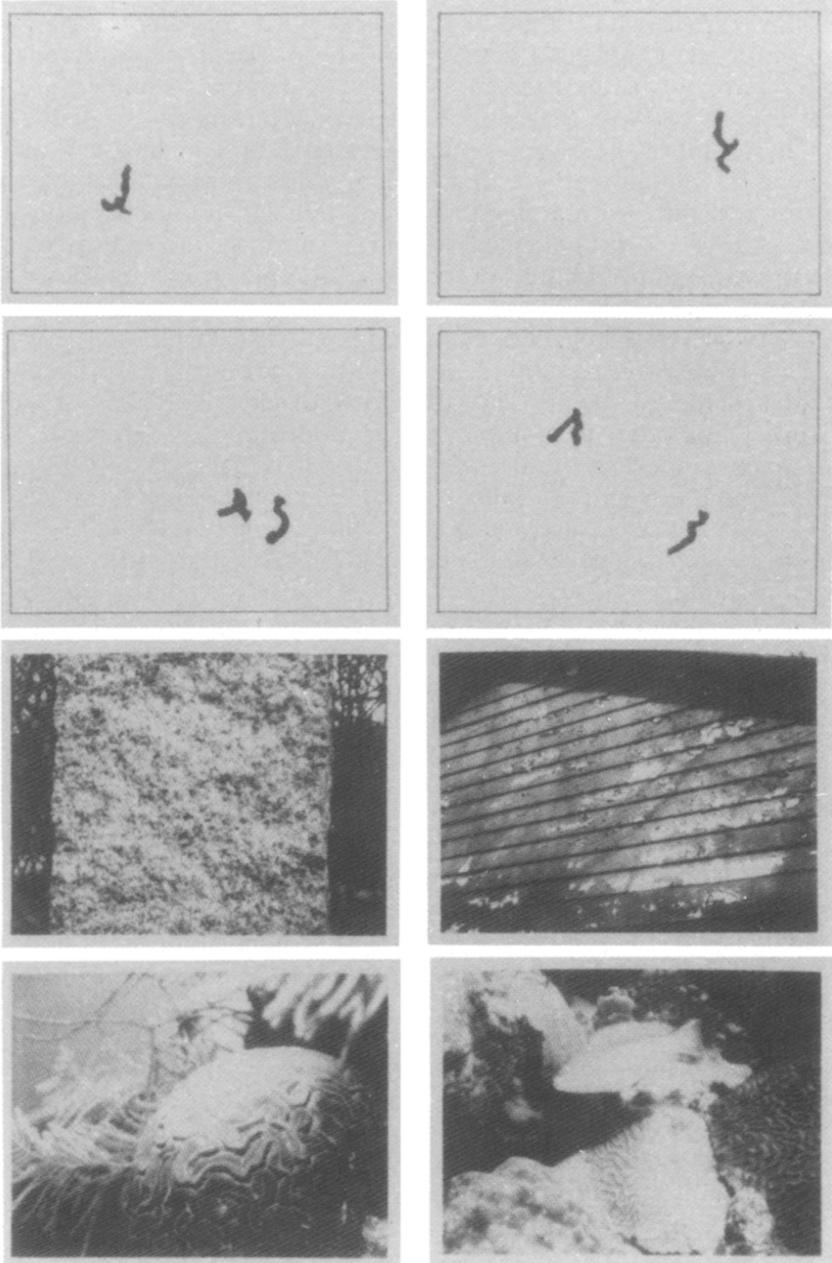
Those interested in categorization as part of human cognitive capacity or as part of the background for human language usually disregard animal data entirely (see, for example: Estes, 1986; Homa, 1984; Smith & Medin, 1981). Even discussions of the categorizations of preverbal infants are likely to omit any consideration of subhuman data (e.g., Younger & Cohen, 1985), as if categorization were a saltation in mental capacity. But categorization is no saltation. It has turned up at every level of the animal kingdom where it has been competently sought.

One reason for looking more carefully at lower levels of categorization is that the continuity of cognitive processes linking humans and other animals is clear and undeniable here. And, as the evidence to be summarized suggests, it is probably at the upper end of this span that animal and human cognitive capacities diverge.

Categorization by rote

Animals can obviously categorize by rote, but there is a question about how powerful a capacity this can be. Using pigeons as subjects, Vaughan and

Figure 1. Examples of the stimuli used by Vaughan and Green (1984) in their experiment on rote categorization by pigeons. (From Vaughan & Greene, 1984.)



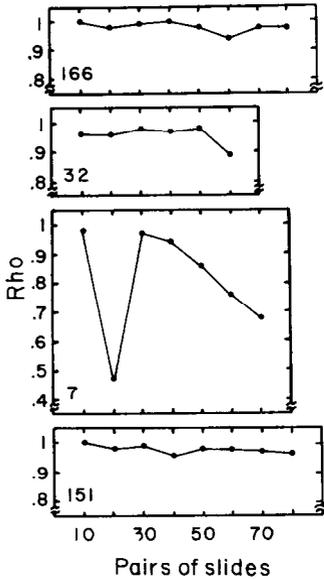
Greene (1984) hoped to examine the sheer size of an animal's storage for a rote category. The general procedure was to assign visual stimuli randomly to positive and negative categories, then to add additional exemplars as a pigeon learned what it had already been given. Pecking a response key was intermittently reinforced with food in the presence of positive stimuli, and not reinforced in the presence of negative stimuli. Stimuli were presented seriatim, for no less than about 20 s each. Further details of the procedure can be found in the original publication (and in Vaughan & Greene, 1983). It need only be noted here that the procedure trains pigeons to peck the key in the presence of exemplars from one category and not to peck in the presence of those from the other. Measures based on the rate of responding in the presence of a stimulus in procedures like this have been shown to provide sensitive and stable estimates of the subject's categorization (Herrnstein, 1979, 1984; Herrnstein, Loveland, & Cable, 1976).

Two kinds of visual stimuli were used for two groups of pigeons – photographs of natural scenes and arbitrary squiggles, examples of which are in Figure 1 – in order to assure at least a modicum of generality to the results. For each kind of stimulus, after an initial population of 10 exemplars (5 positive and 5 negative), pairs were added as the pigeon learned to discriminate what it had already been given. Stimuli were assigned to the positive and negative class by tossing a coin. No superordinate rule could, therefore, reliably identify category membership. As far as the contingency of reinforcement was concerned, this was rote categorization, pure and simple.

The results for squiggles are in Figure 2, for the four pigeons in this part of the study. The measure of discrimination is ρ (Herrnstein et al., 1976), a nonparametric statistic estimating the probability that the average positive exemplar is ranked above the average negative exemplar, using ranked rates of pecking as the measure of performance. With perfect discrimination, ρ is 1.0; with no discrimination, it is .5. For 40 pairs of slides, discrimination reaches a .05 level of statistical significance at a ρ value of .62 (the larger the number of pairs, the closer to .5 is this threshold level of statistical significance). The ρ s in Figure 2, approaching 1.0, signify high levels of discrimination, given the numbers of slide pairs. When the pigeons were sorting 40 or more pairs, the associated p values per session were usually smaller than 10^{-12} .

But what is even more notable than the height of the ordinate in Figure 2 is its general flatness. Two of the pigeons (the other two died from illness or injury) were, by the end of this experiment, sorting with high accuracy 160 exemplars of squiggles, ranking almost every positive above almost every negative. There was no lowering level of performance as the population of slides grew from 10 to 160, nor were the pigeons taking longer to learn new

Figure 2. For four pigeons, rho, an index of discrimination, for increasing numbers of "squiggle" stimuli, which were introduced a pair at a time. The pigeons were required to sort squiggles into two arbitrary categories. (From Vaughan & Greene, 1984.)

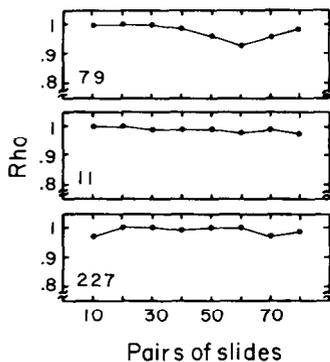


pairs as they were added. In short, it may be said that 160 squiggles do not appear to tax the pigeon's storage capacity for squiggles.

Figure 3 gives the comparable data for the three pigeons working on photographs of natural scenes. Performance is similarly high up to 80 pairs, and similarly gives no sign of taxing storage capacity. If anything, judging from the speed of learning to sort new pairs, scenes were easier to learn than squiggles, despite the much larger amount of visual information ostensibly in them. Stimulus pairs that resembled each other to human observers were harder for the pigeons to learn to sort than pairs that looked different. This simple finding suggests that the pigeons were seeing the photos much as we see them ourselves.

Two pigeons in each procedure were tested for retention of the last 40 pairs of exemplars after more than a year of not having seen them. For the squiggles, the test came 490 days after last seeing any of those particular stimuli; for the scenes, it came after an interval of 629 days. All pigeons discriminated the 80 positive and negative exemplars significantly the first

Figure 3. For three pigeons, ρ for increasing numbers of photographs of natural scenes, which were introduced a pair at a time. The pigeons were required to sort photos into two arbitrary categories. (From Vaughan & Greene, 1984.)

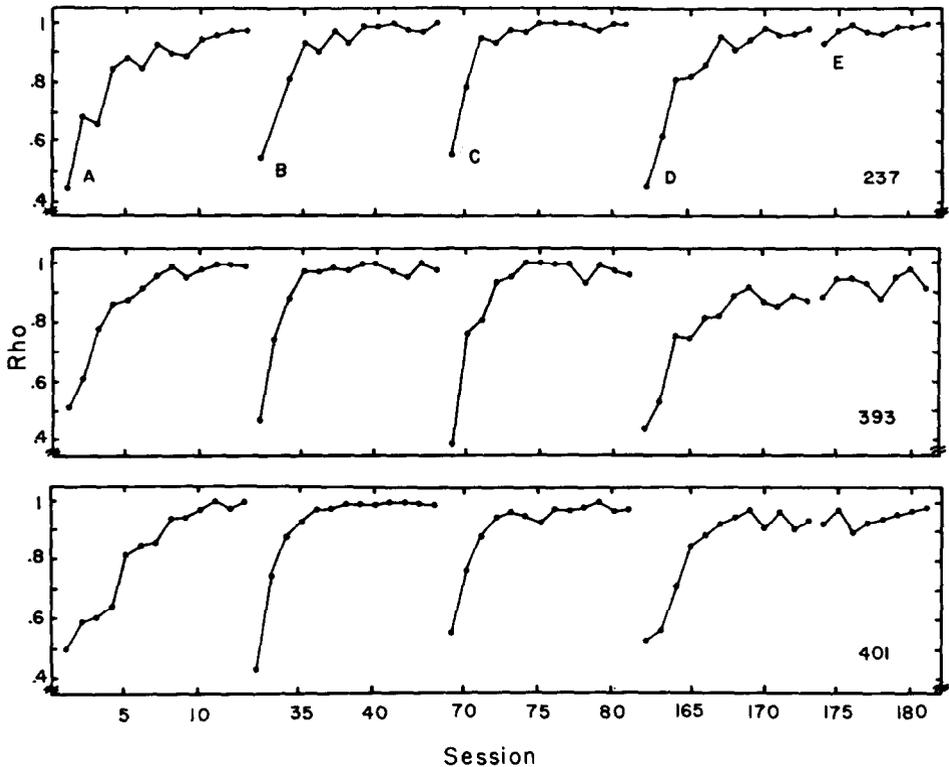


time seen after those intervals. For the very first presentation, values of ρ were down but still highly significant – in the .7 to .8 region – but by the second viewing of the stimuli, the values of ρ almost recovered to their original levels above .9.

Vaughan and Greene (1984) also plumbed the pigeon's rote memory capacity by introducing large numbers of slides at once, not just in pairs as described above. With another group of subjects, 40 pairs of randomly assigned scenes were presented starting with the first training session. The pigeons saw the 80 slides twice per daily session. The acquisition of discrimination is shown in Figure 4. Within 10 sessions, ρ rose to high levels, comparable to those in the earlier experiment. At the session labeled B in Figure 4, a new set of 40 pairs was substituted for the original set. ρ then dropped to approximately .5, signifying no discrimination, as appropriate for a rote sorting task. But then, even more rapidly than for the original set, the pigeons learned to sort the new slides correctly. Yet another set of 40 pairs was used starting at C, and likewise at D. Each set was rapidly learned.

At this point, the pigeons had been trained to sort 320 slides into two arbitrary categories. The eight sessions starting at E show that all those discriminations were simultaneously within the pigeons' reach. On consecutive sessions, the pigeons saw slide sets 4, 3, 2, 1, 4, 3, 2, 1. The points are for the first presentations per session, so that they show retention without benefit of practice within these test sessions. The pigeons were also retested, without additional practice, more than 730 days later. Levels of ρ , on the very first test, were above .7 for all sets of slides.

Figure 4. *New sets of 80 natural scenes, arbitrarily divided into two categories, were introduced at A, B, C, and D. Starting at E, the sets were shown on successive sessions in the order 4, 3, 2, 1, 4, 3, 2, 1. Rising values of rho show the acquisition and maintenance of discrimination for three pigeons. (From Vaughan & Greene, 1984.)*



Numerous control procedures by Vaughan and Greene (1983, 1984) leave little or no plausible doubt that the pigeons were, as the data imply, learning to sort complex visual stimuli on the basis of visual appearance. The rote capacity of this subject is large enough for at least 320 photos of natural scenes or more than 160 arbitrary squiggles. But those numbers may be gross underestimates, for nothing in the learning curves betrays any strain on the pigeon's capacity to impress new exemplars on its memory. The high degree of retention is consistent with this – had Vaughan and Greene been taxing their subjects, a rapid loss of retention would open up storage capacity to

new learning. But, on the contrary, pigeons recalled with high accuracy the first set of 80 slides even after being trained to sort 240 additional ones.

The pigeon's large capacity for arbitrarily classified exemplars is not unique. Shettleworth and Krebs (1986), for example, report that food-storing birds (e.g., marsh tits, black-capped chickadees, and Clark's nutcrackers) remember many more locations of hidden seeds than the number of visual scenes that pigeons have been demonstrated to retain. For estimates of remembered locations in the tens of thousands by Clark's nutcracker, see Vander Wall and Balda (1977; also see Vander Wall, 1982). Kamil's study (1978) of the Hawaiian honeycreeper uncovered evidence of a capacity to keep track visually of which among more than 1,000 flowers the bird had already visited. Human beings, of course, have comparably large capacities for rote categorization (Nickerson, 1965; Shepard, 1967). Perhaps other mammals do too, although not many are as visual as birds and people are, and it is easiest to do the necessary experiments in the visual modality.

Open-ended categories

However large the capacity for arbitrarily classified exemplars, it cannot be large enough for the ordinary demands of the natural environment. Animals usually need to recognize exemplars at levels of aggregation above the individual level. For the recognition of living forms, exemplars of a class are likely to be at the species level or something approximating it. Squirrels recognize the generic acorn, not a particular acorn; likewise cats recognizing mice and vice versa. Inanimate objects are also likely to be generic – *a*, rather than *the*, pool of water, for example. At these levels, variability is the rule, not the exception. Even when individual recognition is appropriate – for example, parents recognizing their young and vice versa, or a particular turn in a brook – the conditions of viewing impose variability. Any natural contingency of reinforcement is therefore likely to be open-ended in the present sense, comprising a virtually limitless set of exemplars from the subject's vantage point. Mere rote will not suffice, unless it also includes some principle of similarity. The inevitability of this demand, and its solution by perceptual similarity, are what suggest that similarity is an evolutionary adaptation.

The evidence is more than anecdotal. Trained, for example, with one set of photographs containing or not containing trees, pigeons generalize to new exemplars (e.g., Herrnstein, 1979). The ability to handle open-ended reinforcement contingencies has frequently been experimentally demonstrated, not just in pigeons but in other animals (e.g., cryptic moths as a category for blue jays [Pietrewicz & Kamil, 1977]; various objects for a parrot [Pepper-

berg, 1981] and for mynahs [Turney, 1982]; Munsell color chips for a chimpanzee [Matsuzawa, 1985a]). The results of a wide variety of experiments on many species leave no doubt of a capacity for categorizing beyond rote memorization (see review in Herrnstein, 1984).

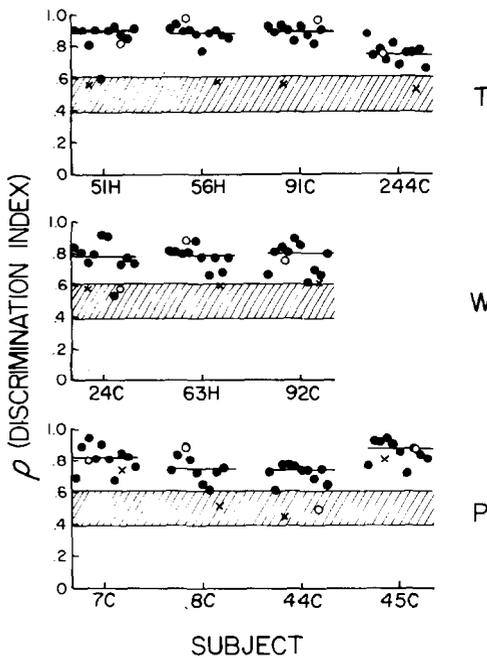
The test of a categorization beyond rote is generalization. Figure 5 reproduces results of generalization tests for pigeons in three experiments (Herrnstein et al., 1976), in which the categories were photographs of trees, bodies of water, and a particular woman in many different settings. The pigeons had been trained by a version of the procedure already described, and they saw differing, but overlapping, sets of photographs once per daily session for several weeks, approximately half containing the relevant object and half not containing it. After this initial training, they were shown 800 new photographs only once, over 10 consecutive sessions, divided equally between exemplars and nonexemplars of the appropriate category. The values of ρ for those generalization tests are plotted as filled points in Figure 5.

Within the cross-hatched area, values of ρ would fall short of conventional statistical significance for discriminating between positives and negatives. Almost all of the filled points are well outside this region. The horizontal lines show medians for individual pigeons, which are all above .7 and most are above .8. The pigeons had generalized whatever it was they learned, and their generalization conformed quite well to what the experimenters considered to be the principle of categorization. The open points and x's summarize attempts to ascertain further whether human experimenters and pigeon subjects were sorting in the same way.

For the regular generalization tests, the slides were drawn at random from a large collection, but the x's came from a session in which pigeons saw 80 stimuli that the experimenters selected as being hard to discriminate. Positive slides had obscure views of trees or water or the person; negative slides had features that could be confused for those objects or were otherwise ambiguous. Nine of the eleven pigeons failed to discriminate on this session; the exceptions were two of the four pigeons in the person experiment. It seemed possible to the experimenters that we humans rely more on faces for identification than our pigeon subjects, but this has not been tested. Except for the two deviant pigeons, this special test strongly suggests that pigeons and humans view these categories similarly.

The other special test, summarized by the open points, was a session using only slides judged easy by the experimenters. Here, no clear pattern emerged. Judging from the values of ρ , some pigeons found this set of slides easy, others found it comparable to the usual generalization test, and for two pigeons (24C and 44C), discrimination failed with these supposedly easy exemplars. The original report attempts to explain these mixed results, but

Figure 5. Filled points give rho values for sessions presenting 80 photographs for the first time, half containing, and half not containing, the relevant object (T = trees; W = water; P = person), for 11 pigeons. Open points for sessions presenting photos judged easy to categorize by the experimenters; x's for sessions presenting difficult photos. Inside the cross-hatched region, rho is not significantly different from .5, which would indicate no discrimination. (From Herrnstein, Loveland, & Cable, 1976.)



whatever the explanation, from this test, we could not be encouraged to infer complete equivalence of the categorizations of people and pigeons. Taking the two tests together, we may tentatively conclude that the experimenters and pigeons were working with overlapping, but not precisely equivalent, categories. The lack of equivalence should not be surprising, not just because humans and pigeons are different species, but because their reinforcement histories are so different with respect to the particular categories and to the reading of photographs, and because the methods used to assess the categories were so different.

It may be unwise to equate similarity with mere confusability or imprecision in perception, as is often done. The dimensions of similarity can only be

inferred from behavior, but the pattern of data from many experiments makes it clear that something more dynamic than mere error in discriminative precision is afoot. There is a long tradition in formal theories of similarity (e.g., Restle, 1959; Tversky, 1977) according to which two items are similar as a decreasing function of the elements (or attributes) they do not share and as an increasing function of the elements they have in common. This insight may be formalized in many ways, as long as it expresses that similarity depends upon context. The same pair of printed capital A's in different typefaces may look similar among an assortment of different letters of the alphabet and different among a collection of varying upper-case A's. Attributes that produce similarity in the first set give way to attributes that produce difference in the second. Theories assume something like a periodic table, or alphabet, of attributes, which are then compounded or spelled into the categories at hand. Promising as this approach has been, it has insufficiently captured in any formal model yet articulated how plastic and diverse the attributes themselves appear to be.

When adequately challenged, subjects have an uncanny knack for finding surprising attributes to help them distinguish between categories. To uncover this knack, it is essential that it be profitable for the organism to resort to it. Three studies should make the point. In an experiment using blue jays as subjects, Real, Iannazzi, Kamil, and Heinrich (1984) took advantage of the differing leaf-damaging patterns made by different species of caterpillars. The authors distinguished between a particular pair of caterpillar species by saying that one damages leaves "neatly," whereas the other does so "messily." This difference itself is probably an evolutionary adaptation, inasmuch as the neat eaters are palatable and the messy ones are not. Neatness, by disguising the traces on a leaf of having been nibbled on, protects the palatable caterpillar who did the nibbling.

Blue jays were trained to behave in two different, but arbitrary, ways in the presence of silhouettes of leaves nibbled by the two species. Only one leaf of each type was used in training, yet the blue jays generalized to other exemplars readily, responding in one way to leaves attacked by one species and in the other way to those attacked by the other. The birds had, from only one example, learned to sort leaves according to the nibbling patterns of the caterpillars. Where or how much or precisely how the nibbling was distributed was so diverse that, according to the authors, "no simple rule of geometry ... could lead to correct categorization of all of the novel stimuli" (Real et al., 1984, p. 208). Whatever the geometrical algorithms that could discriminate between the classes, the blue jays proved themselves to be capable of finding just the right stimulus attributes to master an open-ended categorization based on the nibbling patterns of caterpillars.

The two other experiments to be described here seem to me to provide even stronger evidence of the adaptability of open-ended categorization. Vaughan and Greene (1984) showed that pigeons can learn to sort a set of 80 slides into two categories, but the sorting rule was conditional on mirror orientation. The 80 slides were arbitrarily divided into two classes, A and B. No single class of objects, such as trees, or any other perceptual feature, characterized either category. The pigeons learned to sort the two arbitrary categories, but then all 80 slides were mirror-reversed. With mirror reversal, all the reinforcement contingencies were also reversed: If Class A was positive in the original orientation, it was negative when mirror-reversed, and vice versa for Class B.

Initially, because pigeons readily generalize to mirror reversals, they responded inappropriately: The pigeons continued to respond to exemplars of Class A and not to those of Class B, even when reversed in orientation. But they learned the new rules within several sessions. When they were properly responding to the reversed categories, the contingencies were again reversed, along with another mirror reversal. The slides were now in their original orientation, with the original contingencies of reinforcement. Learning was more rapid, and, after it was completed, once again orientation and contingencies were reversed. This pattern of reversing orientation and contingencies continued until the pigeons learned to respond appropriately to each and every slide, depending upon its orientation. The rule that eventually came to dictate behavior was: If a slide was positive as originally oriented, then it was negative when mirror-reversed, and vice versa. Without the peculiar training of this procedure, pigeons generalize spontaneously and fully to mirror reversals, but they learned to use something arising in reversal as a cue to categorization. Orientation, or some by-product of orientation, emerged as a controlling attribute when it paid.³

The final experiment used pigeons who had homed either around Ithaca, NY, or Lincoln, MA, or who had not homed at all (Gray, 1987). The stimuli were aerial photographs, landscape scenes from 15 to 25 square mile regions in the vicinity of the two homing lofts, and of a comparably rural area mostly near Georgetown, NY, where none of the subjects had flown. All three regions seemed similar in topography, vegetation, and population density; Ithaca and Georgetown seemed particularly similar. The pigeons were trained to respond, in a standard laboratory procedure, in the presence of photo-

³In this experiment, in which orientation was not consistently associated with reinforcement or nonreinforcement (it was associated one way with Class A slides and the other way with Class B slides), nothing decisive could have been learned from testing the subjects with novel slides in one orientation or the other. Although the results might have raised some interesting questions, no such test was attempted.

graphs of the Ithaca and Lincoln locales, but not of the Georgetown locale. Whatever a pigeon's past history, it learned quickly and well to discriminate between scenes from Ithaca or Lincoln and Georgetown. Having homed in an area did not seem to confer any special advantage in recognizing it in photos. All pigeons also generalized successfully to photographs taken in a locale but not used in training, or not overlapping with those used in training. This finding suggests that pigeons recognize scenes from a locale as such, based solely on samples seen in photographs. To the experimenter or her associates, including the present author, who have studied those photographs, the cues to locale remain obscure.

In these three studies, we may presume visual attributes, without having anticipated them on the basis of other findings with blue jays or homing pigeons. The attributes for leaf damage, locale, or orientation may be behaviorally invisible until the contingencies of reinforcement highlight them. The results point to a large theoretical gap, not in a formula for subdividing or aggregating attribute sets, but in an account of the dynamic character of the attribute sets themselves.

Concepts

Similarity was the distinguishing element as we moved from rote to open-ended categories. The distinguishing element in the further step to concepts is harder to say or to grasp. Indeed, it may safely be assumed that some readers may be, and some will remain, unconvinced that this further step is necessary. Until I read the relevant essay by Stephen Lea (1984), I was unconvinced myself. What follows is a formulation inspired by and similar to, but, in detail and in substantiation, differing from, Lea's.

Let us start by assuming that the mythical grandmother cell (e.g., Barlow, 1972) exists. Someplace in the brain, a neuron becomes active whenever granny appears in any of her manifestations. Such a cell would be at the apex of a hierarchy of neural analyzers of increasing generality across such variables in the proximal stimulus as size, location, luminance level, color, perspective transformations, and so forth, and increasing specificity for invariance in the distal, objective world, ultimately delimited to granny herself. The grandmother cell could itself establish connections with other functional loci – sensory, motor, or motivational. The organism may find the cell's activation reinforcing or punishing. Its activation may evoke activity in the system controlling speech – for example, the utterance “granny.” The point is simply that the cell operates as a functional entity in the nervous system, entering into relationships available to functional entities.

The notion of a concept is like the notion of a grandmother cell, with one vital qualification. A concept need not be mediated by a neuron or any particular set of neurons. It need not even be localized in neurons, rather than, say, in synaptic neurochemistry or in more global patterns of activity in the nervous system. As far as the *definition* of a concept within the present scheme is concerned, behavior, not neurophysiology, is central. However, if a case for concepts can be made behaviorally, then finding its neurophysiological basis becomes a proper challenge.

A concept is, to begin with, a basis for categorization, a differential response of the organism to a set of stimuli. The set may be limited or open-ended, as in the two levels of categorization already discussed. Beyond that, however, the effects of contingencies applied to members of the set propagate to other members more than can be accounted for by the similarities among members of the set. In the ideal case, they propagate to all members of the set, without regard to similarity. Because similarities for a given organism under particular circumstances can themselves only be inferred from behavior, we may expect difficulty in demonstrating concepts, as distinguished from the lower levels of categorization. It is also the case, however, that what seems to be merely open-ended categorization, may turn out to be, when appropriately tested, a case of conceptualization.

In vocabularies that have fallen into disuse, “mediated generalization” and “mediated association” resemble what is here being called a concept (see, for example, Osgood, 1953, p. 359f, and Underwood, 1966, p. 533f, respectively). The terms were used to describe transfer of conditioning across stimuli too different to be attributed to common features or attributes. The theoretical notion was that the stimuli may have had no stimulus elements in common, but, presumably because of past conditioning, they evoked overlapping responses. The overlapping responses provided the bridge for transfer. For example, imagine conditioning the eyeblink response of a human parent to the word “baby.” Transfer would then be expected to the word “cradle.” Baby and cradle fall into a common category, but not because of similarity as ordinarily defined, but because of a connection at some other level of aggregation.⁴ Unlike the earlier writers, I make no theoretical assumptions about the bridge for the mediated generalization. The issue here is simply whether analogous examples of categorization can be convincingly shown for nonverbal animals.

For mediated generalization, the link between members of a set is presumed to be based on learning. For grandmother cells and the like, it has, at

⁴In words, the concept would be something like “pertaining to my baby.”

least by some workers at some times, been supposed that they are substantially hard-wired (e.g., Hubel & Wiesel, 1962). The present scheme is agnostic on this point. It is assumed that concepts may be either learned or more or less hard-wired.

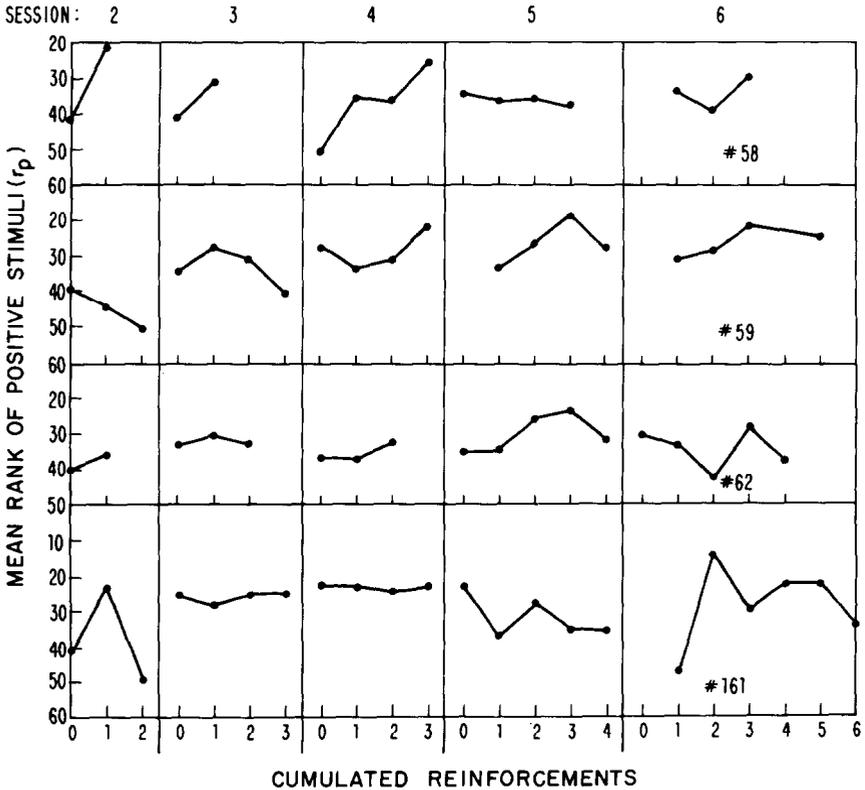
Suggestive evidence for concepts in pigeons can be found in an experiment already cited as showing an open-ended category (Herrnstein, 1979). Pigeons learned to sort slides containing trees from those not containing trees. The training procedure consisted of showing 40 exemplars and 40 nonexemplars once per daily session. On successive sessions in this particular experiment, the pigeons saw the same 80 stimuli in different orders. Because the same stimuli were used daily, it is possible to draw a plausible inference about whether the pigeons were sorting on the basis of exemplars or a concept.

The reinforcement schedule for responding in the presence of exemplars was so lean that as the discrimination was being formed, exemplars were associated with varying total numbers of reinforcements summing across sessions, from zero on up. Figure 6 plots, along the x-axis, the total number of reinforcements associated with given exemplars. The y-axis plots the mean rank of the exemplars with the indicated numbers of reinforcements. Mean rank is a measure of discrimination; it should "rise" from 40.5, when there is no discrimination, toward 20.5, when all 40 exemplars are ranked above all 40 nonexemplars. The figure shows sessions 2 to 6, which is when the animals were learning the sorting task.

If the pigeons were sorting open-endedly by exemplars, it would be reasonable to expect the lines in Figure 6 to have positive slope, indicating that the more reinforcement an exemplar had in its history, the better it was discriminated. Pigeon 58 had lines of positive slope up to Session 4, but this was the slowest learner of the four (see the original report for details). More reinforced exemplars were discriminated better. The other pigeons had flat lines from the start (as did Pigeon 58 after the fourth session). The level of performance rose for all exemplars, without regard to reinforcements in their history. Even exemplars with no reinforcements were rising with the set.

It is plausible to infer that, when the lines were flat down to zero on the x-axis, sorting is on the basis of something more general than exemplars and similarity to those exemplars. Were such results obtained from a human subject, we would have no hesitation in invoking the concept called "tree." On the other hand, it may be argued that the unreinforced exemplars were, by virtue of how pigeons average exemplars (assuming that is what they do), no less similar to the reinforced exemplars than the reinforced exemplars were to themselves. But this argument would be indefensible on its face, because, in this procedure, which stimuli were unreinforced was an accidental and variable outcome of the reinforcement and stimulus presentation

Figure 6. Photographs containing trees (i.e., positive stimuli) are grouped according to the total number of reinforcements received in their presence, counting from the beginning of the experiment to the session number indicated at the top of the figure. The mean rank of positive stimuli (r_p) is a measure of discrimination, which is at 40.5 at no discrimination and advances to lower values as discrimination forms. Four pigeons are represented. (From Herrnstein, 1979.)

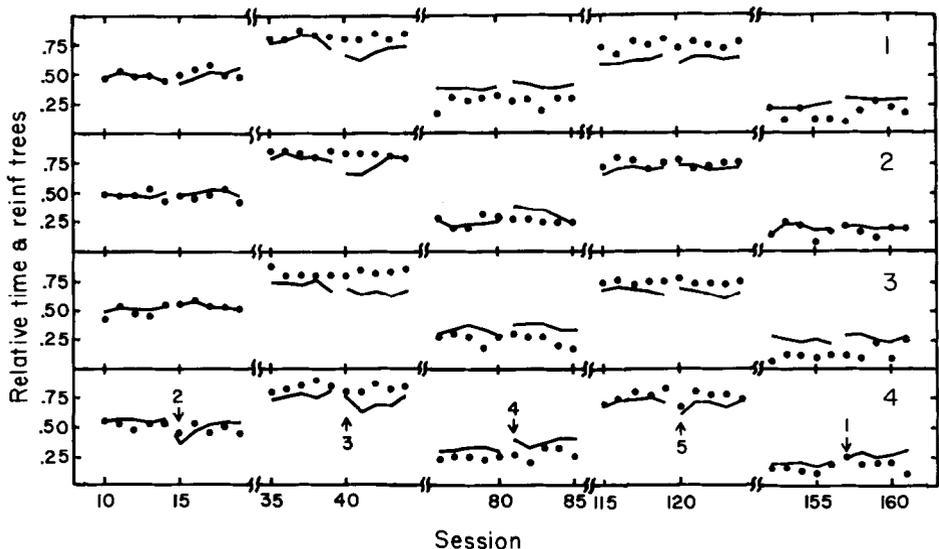


schedules and, more relevantly, differed for each subject. It is highly improbable that those differing sets of exemplars appropriate to the averaging hypothesis for each subject coincided, by chance, with the unreinforced exemplars for each subject. The effects of reinforcing a few exemplars of trees propagated with no decrement to an open-ended set of exemplars of trees, independent of the number of reinforcements per exemplar. A tree-like concept may be denied to these pigeons, but not plausibly.

In another study, pigeons were again required to discriminate between photographs containing trees and those not containing any (Vaughan & Herrnstein, 1987). Responding was reinforced in the presence of all slides, but different schedules provided reinforcements for responding in the presence of trees and nontrees. Reinforcement sometimes came at a higher rate in the presence of trees than in the presence of nontrees; at other times, vice versa. Under these conditions, subjects approximately obey the "matching law," responding to the two categories in the same ratio as of the respective rates of reinforcement (Herrnstein, 1970).

During the course of the experiment, the ratio of rates of reinforcement was changed four times, and, each time, the pigeons suitably adjusted the ratio of their rates of responding. In addition, during the course of the experiment, five sets of 80 different photographs were shown, each one including 40 containing, and 40 not containing, trees. A summary of the results is shown in Figure 7.

Figure 7. *Proportion of reinforcements received (points) and time spent (lines) in the presence of photographic slides of trees, for four pigeons. According to the matching law, points and lines should be superimposed. The initial sessions used Slide Tray 1, which was replaced as indicated by Slide Tray 2, and so on, until Slide Tray 1 was restored at the 157th session. Each slide tray contained different slides, of which just half contained trees. (From Vaughan & Herrnstein, 1987.)*



The solid lines show the proportion of time spent in the presence of the tree slides, for five consecutive sessions before and after the stimuli were changed. The subjects themselves controlled stimulus duration, so this is an appropriate measure of behavior. The points track relative rate of reinforcement. The approximate convergence of points and lines exemplifies the matching law. The numbers along the lines show where sets of slides were replaced; for example, at Session 15, Slide Tray 1 was replaced by Slide Tray 2, and so on. The replacements were at no point coincident with a change in the reinforcement schedule, and vice versa. At each change in slides, the pigeons continued responding in approximate accordance with the matching law. This continuity of behavior demonstrates generalization from one set of slides to another, a case of at least open-ended categorization and, possibly, conceptualization.

The results of the final replacement, labeled 1, suggest a concept, which is to say, something more than the sorting of particular exemplars plus extensions owing to similarity. Here, Slide Set 1 replaced Slide Set 5. Responding generalized: The approximately .25 proportion of time in the presence of trees transferred from Slide Set 5 to Slide Set 1. However, the pigeons had seen Slide Set 1 at the beginning of the experiment, when they had been spending about .5 of their time in the presence of tree slides (see Figure 7). At the final replacement, the pigeons generalized the behavior appropriate to the category of trees as a whole rather than the behavior previously conditioned to the particular exemplars in Slide Set 1. The current contingency of reinforcement had propagated to the exemplars in Slide Set 1, displacing the effects of the contingency of reinforcement in force when Slide Set 1 was last seen, which satisfies the present definition of a concept.

A different approach to concepts is exemplified in an experiment by Vaughan (1988). Pigeons were shown the same 40 tree-containing slides in each session, always in a different, random order. The slides were divided into two arbitrary categories, one positive (reinforcement in their presence), the other, negative (no reinforcement). Different divisions of the 40 slides were used for different pigeons, so as to eliminate the possibility that some unknown perceptual feature could signal category membership. After the pigeons learned to sort the slides, reinforcement contingencies were reversed – positives were now negative and vice versa. When this new task was mastered, the categories were re-reversed, restoring the original contingencies. The reversals continued until the pigeons learned to detect, from the first few slides in a session, whether the contingencies were as they were originally or reversed. The pigeons had, in effect, created a file for each category, and could tell by sampling a few slides (less than a half dozen) whether the file had a positive or negative valence on any particular session.

No principle of similarity as perceptual proximity can account for these results, since the slides in either category were no more similar to each other than to the slides in the other category. Nor does the usual notion of a conditional discrimination apply, for no additional stimulus, signalling the conditional relationship, was provided. The contingency of reinforcement at the beginning of a session was extrapolated by the pigeon to all the slides in the tray. Vaughan's experiment demonstrates an arbitrary concept, with small, rather than open-ended, membership, namely the 20 slides in each category.

The operational distinction between concepts and open-ended categorization is that something besides similarity gradients permits new exemplars to be categorized. This "something" is a representation within the organism at a more inclusive level than individual exemplars and their similarity domains. Since similarity itself can be inferred only from behavior, distinguishing between these alternatives may seem a futile undertaking or worse. However, human beings evidently conceptualize in the present sense in the very act of using language. Language is a sufficient condition to permit the inferring of concepts, but it is not also a necessary one. The evidence suggests that non-linguistic creatures conceptualize as defined here.

The difficulty of distinguishing concepts and categories cuts both ways. Some people may be disposed to argue that the examples presented here of concepts are more parsimoniously interpreted as mere categories. On the other hand, the inverse argument is also possible. What looks like simple categorization based on similarities may, if suitably analyzed, prove to be based on concepts.

Abstract relations

Concepts may stand in relation to each other. Members of Class A may be bigger, longer, faster than, above, to the right of, prior to, or just different from, members of Class B. Relations may be quantitative: Members of A may have half the area of members of B. They may be logically dependent on other conditions: There is leaking water in the Red Line stations only after it rains a certain amount. The psychological questions are whether, and to what degree, behavior can be brought under the control of a relation between concepts, what I am here calling an abstract relation. It is *abstract* because it should be generalized to all exemplars of the relevant classes, even those exemplars not previously experienced in the given context.

The prototypical laboratory demonstration of an abstract relation is a matching or oddity procedure. In their close conceptual proximity to simple

category membership, matching and oddity may be considered among the most elementary abstract relations, yet they are a step beyond. A subject is reinforced for choosing, among a set of alternatives, that choice which matches, or, in the case of oddity, does not match, a standard in some respect. The subject may, for example, be choosing the color match of the standard. After training with, say, red, blue, and green standards, the subject is tested with a yellow standard. If it chooses correctly, it has given evidence of an abstract relation, one we may equate with "the same color as." In the oddity version, the subject would be required to choose the one color that differs among a set of three or more color alternatives. It would again be tested with new sets of colors, not previously used in training.

If the experiments are well constructed, then the generalization of training cannot be accounted for by anything other than an abstract relation itself – such other accounts, for example, as unwitting stimulus similarities between training and test stimuli. Even when the obvious plausible alternative explanations of the data are excluded, an inescapable ambiguity remains about the scope of the relation. From what was said above about the color experiment, it would have been just as correct, logically, to infer the abstract relation "the same as," or "the same visual stimulus as," rather than "the same color as." A narrower construction was also possible – "the same color as, except for yellow." Each such relation dictates its own implications for tests of generality.

The subject's task in an experiment on abstract relations is to draw an inductive inference. This is also the case with any categorization problem in which the subject must extrapolate to new exemplars. In testing for any such categorization, we are inexorably limited by the indeterminacy of induction. If, for example, the subject is using an abstract relation to categorize stimuli, then only the subject can, by its behavior, reveal its scope. An animal may fail our test of generality, yet still be using an abstract relation consistent with the evidence it had to go on.

Many experiments have explored matching, oddity, and other logically analogous procedures, with nonhuman subjects, since the pathbreaking work of Harry Harlow and his associates on primates (Harlow, 1949; Meyer & Harlow, 1949). Some of the research and controversy surrounding chimpanzee language verges on the question of the capacity to learn abstract relations and the degree to which that capacity is the evolutionary precursor of human linguistic ability (e.g., Premack, 1976, 1986; Ristau & Robbins, 1982; Rumbaugh & Pate, 1984; Terrace, 1979, 1985). The simple declarative sentence itself is the very embodiment of an abstract relation, with its subject and object linked by a verb. Even the most concrete example (e.g., "A dog bit a man") conforms to the paradigm – two concepts in a particular relation. So

it is no surprise that here is where categorization and language converge, and where the first clear traces of differentiation among species have been observed.

In one recent experiment, Cebus monkeys learned the identity relation over sets of visual stimuli (D'Amato, Salmon, & Colombo, 1985). They were first trained to match one or the other of two simple stimuli drawn from a small collection of such stimuli – dots, triangles, vertical lines, and so forth. On any given trial, one or the other of the stimuli was the standard, and both of the pair were presented simultaneously as alternatives. Reinforcement was earned by picking the matching alternative. After learning this task, the monkey was tested with a new pair of stimuli as standards and comparisons, to see if identity as such controlled choice, or at least if learning was much more rapid than with the original pair. Of eight monkeys in the sample, four met a criterion for transfer after learning with the first pair; two, after learning the identity problem with two pairs; and the seventh, after three pairs. The eighth monkey was dropped as inept at matching early in the experiment. Of four monkeys who successfully transferred matching with steady stimuli drawn from the original set, none transferred when tested with flickering stimuli.

Pigeons in a similar experiment fell even shorter of generality in abstract matching (Holmes, 1979). Indeed, the author concluded that what little evidence of the identity relation there was in his data could be discounted as owing to a nonspecific improvement in performance, having nothing to do with matching as such. Other studies have found better evidence for abstract relations by pigeons. But they have usually also found that the abstract relation is narrow and easily disrupted, which may explain why the evidence as a whole is as mixed as it is (for some indication of the lack of consensus, see Carter & Werner, 1978; Pisacreta, Lefave, Lesneski, & Potter, 1985; Zentall & Hogan, 1978).

Pisacreta and his associates have developed procedures that guide pigeons to use the abstract relations of matching and oddity more reliably than is typical (Pisacreta et al., 1985; Pisacreta, Redwood, & Witt, 1984). In these experiments, the stimuli are usually complex – drawings or photographs of faces or animals, for example – more than two alternatives to choose among are often presented, and the alternatives are presented serially, so that the pigeon is, as it were, allowed to make its choice about alternatives one at a time. Which, if any, of these procedural innovations are responsible for the success, I cannot say. Delius and his associates have also been relatively successful in their matching and oddity experiments on pigeons (Delius & Nowak, 1982; Holland & Delius, 1982; Lombardi, Fachinelli, & Delius, 1984), using novel procedures of their own. Urcuioli has shown, in his exper-

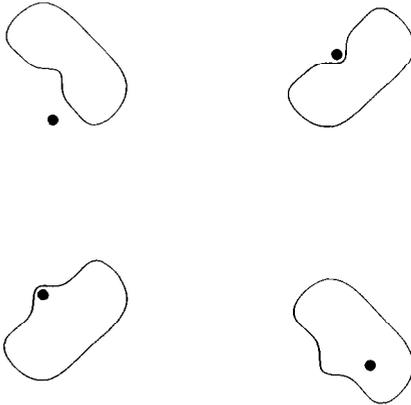
iments (Urcuioli, 1985), that pigeons do better with the identity relation when they are trained to behave differently for the different standard stimuli they are to match, but this procedural detail does not appear to explain the successes of Pisacreta and Delius and their associates. The point is that procedural details appear to be far more critical when pigeons are being asked to use abstract relations than when monkeys or apes are, or when pigeons are asked to categorize at a lower level.

Questions of psychological capacity are tricky (Herrnstein, 1985). It has been observed that a procedure that evokes moderately successful use of the identity relation by monkeys fails to do so with pigeons (Wright, Santiago, Urcuioli, & Sands, 1983). How shall this be interpreted? The authors of the study suggest, guardedly, that monkeys can master the abstraction, but that pigeons cannot. An even more guarded conclusion would be that, in their experiments, monkeys did and pigeons did not. The existential question of capability, a negative answer to which verges on the assertion of a universal negative, should not be answered hastily. Pisacreta's, Delius's, and Urcuioli's data, among others, illustrate why it is prudent to be so guarded about questions of capacity, for they show pigeons to be capable of categorizing according to an abstract relation. The capacity may be fragile, but it is not wholly absent.

A similar conclusion is suggested by the results of an experiment testing whether pigeons can learn to use the abstract relation of "inside–outside" (Herrnstein, Vaughan, Mumford, & Kosslyn, 1989). Inspired by Ullman's discussion of "visual routines" (Ullman, 1984), we tested pigeons with stimuli consisting of a closed curve and a dot, which the pigeons were required to sort according to whether the dot was inside or outside the curve. The stimuli were constructed so as to permit no local cues to insiderness: In the vicinity of the dot, the curve could be either convex or concave, whether the dot was inside or outside. Several of the stimuli are shown in Figure 8. Human observers find the discrimination trivially easy. But the pigeons failed to learn to categorize even a set of 40 dot-inside and 40 dot-outside slides, shown repeatedly in a changing random order, let alone to generalize to other exemplars.

The pigeons were then subjected to a "tutorial" for inside–outside. The insides of the white closed curves were made bright red, while the backgrounds were black. Now the task no longer relied on an abstract relation. The pigeon could sort according to whether the white dot was on a red or black field, a task easily learned. Next, the red was darkened – still discriminably different, to the experimenters and presumably also to the pigeons, from the black background. The pigeons continued to categorize accurately. Then the dark red was replaced by the same black as the background, so the

Figure 8. *Four typical stimuli used in an experiment with pigeons, requiring categorization according to the abstract relation of insideness. (From Herrnstein, Vaughan, Mumford, & Kosslyn, 1989.)*



task was, once again, a matter of the abstract relation of inside versus outside. The pigeons now sorted accurately. They also transferred categorization to new exemplars, so long as the new closed curves were not too radically different from what they were familiar with and did not have invaginations or protuberances that were extreme. An abstract relation, of limited scope, was being used for categorization.

Abstract relations and their derivatives may increase in complexity without bound. A step up the ladder of complexity are the cardinal numbers. All sets of, say, five objects are in the relation of identity for number. The number itself is the concept whose members are sets of objects of the given cardinality, five in this case. With considerable effort and ingenuity, Pepperberg trained a parrot to utter the correct number for sets of two to six objects (1987). Alex, the parrot, generalized to new sets of objects, and even to heterogeneous collections of objects. Two chimpanzees were similarly trained by Rumbaugh, Savage-Rumbaugh, and Hegel to “sum” (albeit not vocally) across objects in an experiment involving the abstract relation, greater-than (Rumbaugh, Savage-Rumbaugh, & Hegel, 1987). There are many reports of counting in animals, but the famous tale of Hans, the clever but noncounting horse, dictates caution (Pfungst, 1911). These experiments are, however, convincing, and so is a similar one by Matsuzawa, using a single chimpanzee (1985b).

Yet even Alex, the parrot, and the various chimps fall far short of display-

ing a set of numerical concepts robust yet flexible enough to be woven into the logical structures of arithmetic and mathematics. No one has yet, to my knowledge, even attempted to train an animal to solve a linear equation in one variable. Even suggesting the possibility seems fatuous. Where the abstract relations among concepts are themselves concatenated into more highly derived conceptual structures, human capacities diverge sharply from the animal precursors. From this perspective, the difference between our closest primate relatives and many other animals, including birds as well as mammals, seems far smaller than the difference between our species and all others.

Conclusions

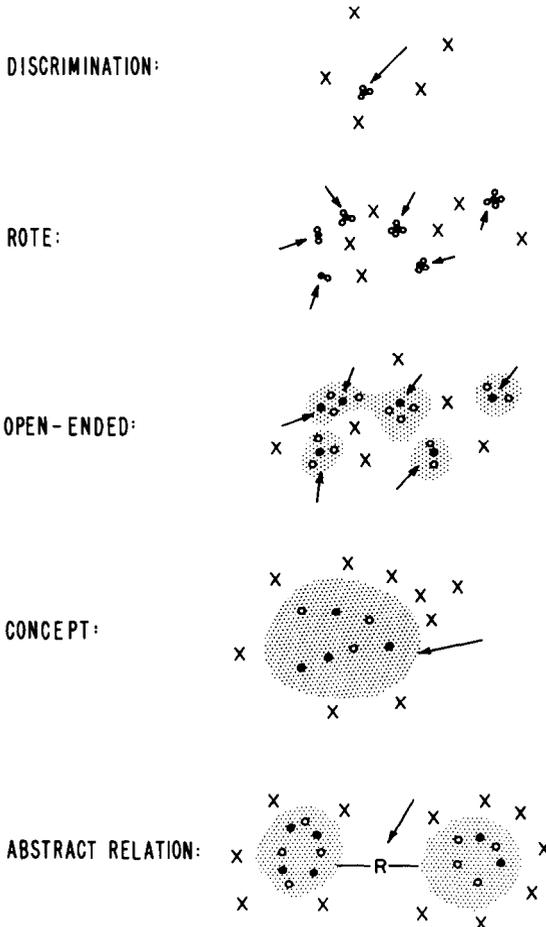
Some, but not all, of this story of levels of stimulus control, as well as some of its gaps, are illustrated in Figure 9. The arrows point to what may be called the loci of stimulus control, corresponding to the five levels of categorization, from discrimination to abstract relations. The arrows indicate how the objects to be categorized are linked to, or are in control of, behavior. Filled circles are exemplars previously experienced, which is to say, exemplars confirmed by past consequences of behavior; open circles, exemplars to which generalization appropriately takes place, given the contingencies of reinforcement. At the two lowest levels – discrimination and rote categorization – the loci of control comprise descriptions of the familiar exemplars themselves, and there is only as much generalization as would be allowed by psychophysical indiscriminability.

At the next level, open-ended categorization, the exemplars still constitute the loci of control, but now generalization extends further from those points, owing to similarity. A category approximates a continuous region in what may be called the animal's subjective quality space, rather than a finite set of points, because of the spread of similarity. The spread is probabilistic, in the sense that near exemplars are more likely to be reached by generalization than remote ones.

Earlier, I noted that animals are remarkably good at finding whatever attributes of a set of stimuli serve their purposes in relation to the contingencies of reinforcement. The spread surrounding exemplars is thus not fixed in the animal's perceptual apparatus, but an adaptation to a given set of reinforced exemplars. The adaptability must, however, be limited: first, by the information provided by the animal's sensory receptors, and, second, by any inherent similarity gradients or other organismic predispositions ("prepared" associations and the like, see Herrnstein, 1977; Shettleworth, 1972). An ani-

mal cannot, for example, make use of potentially useful information from regions of the electromagnetic spectrum where it is blind, even if this information could predict reinforcement. Likewise, similarities or associative biases built into the organism may, in some circumstances, impede adaptation, as useful as they are in other circumstances.

Figure 9. *Levels of categorization. Arrows point to the loci of stimulus control; filled circles to confirmed exemplars; open circles to exemplars to which generalization is appropriate, given the contingencies of reinforcement, and x's to exemplars to which generalization is not appropriate.*



At the concept level, the locus of control no longer resides with the exemplars, but at a more inclusive level, the conceptual grandmother cell discussed earlier. The concept is inferred from, or merely activated by, experienced exemplars. Other exemplars, beyond, or orthogonal to, those directly linked by gradients of similarity, are then responded to appropriately. The category has become truly a region in a subjective space, but probably not a sharply defined one and not one of which the dimensions need be constrained by similarity. The concept's "perimeter" may be probabilistic, as human concepts, with their prototypes and peripheral exemplars, have been shown to be (Rosch, 1973; Smith & Medin, 1981). Robins are better examples of birds than penguins, for example. Similar structures within concepts may be inferred from data on nonverbal categorizers (e.g., see Herrnstein & de Villiers, 1980, for what constitutes good and bad examples of fish for pigeons). As Vaughan's experiment (1988) on arbitrary concepts shows, a concept may comprise members not linked by similarity or anything else, except their common membership. Next to nothing general is known about how the perimeter of a concept is set, but we may guess that it depends on interaction between contingencies of reinforcement as they have been experienced or on perceptual or motivational predispositions.

The fifth level of categorization, abstract relations, covers a multitude of complexity. At its lowest level, it is a simple relationship between concepts, such as identity or difference in some respect. Were this article going on to consider human behavior *per se*, especially language, this level would demand further elaboration. However, even for the most elementary abstract relations (oddity, identity, and the like), the evidence reveals substantial interspecies variation and also a large gap between humans and the closest primates.

We may surmise that the survival demands on animal life account for the ubiquity of at least the first three levels, up to open-ended categorization. Objects in nature, set into the classes that are behaviorally relevant in the life of animals, vary. Sometimes a perceptual system evolves to counteract the natural variation. Ticks need mammalian blood to survive, so they must be able to recognize exemplars of this class, which is so diverse from our point of view. Mammals are not diverse for ticks, which, trading perceptual richness for efficiency, have evolved to detect their hosts by their warmth and by a volatile compound emanating from all mammalian flesh (von Uexkull, 1934, see Herrnstein, 1982). But if the natural variation of objects is echoed in a perceptual system, then open-ended categorization is required.

The next level, concepts, is a less irreversible approach to efficiency than the tick's. The economy of concepts is illustrated in Figure 9 by the shift from many loci of control at Level 3 to just one at Level 4. Concepts are especially

adaptive for creatures that learn, for it enables the implications of a change in the consequences of behavior, or in the class of stimuli with which the consequences are associated, to be rapidly propagated across the members of the class. The rapid propagation is economical, but may sometimes be wrong. As noted earlier, any categorization that extrapolates from past experience is heir to the risks of inductive inference.

Efficiency is promoted further by the step up to abstract relations, for a small amount of logic can displace a large amount of description. Each abstract relation can be considered a substitute for a new concept, of which the exemplars would be the related sets of objects in the abstract relation. Instead of the abstract relation, "this chip is the same color as that chip," a concept could constitute the class of pairs of color chips of the same color. The classes of triplets or quadruplets of like-colored chips, and so on, would each require a concept of their own. But describing pairs, or triplets, or quadruplets, of chips of the same color, per se, is needless mental work if the organism can use the abstract identity relation applied to the concept of a chip, taken in sets of any size.

It may be needless in another sense as well. Abstract relations often express transient states of the animal's environment. The contingency of reinforcement may be the identity of the colors on one occasion, but a difference in color on another. The relation, "acorn on a stump," may be of less enduring interest to a squirrel than the concepts of either acorn or stump. To the extent that abstract relations spare organisms the effort of constructing concepts they are unlikely to resort to thereafter, they are a clear benefit.

Beneficial as they are, abstract relations differentiate most sharply among animals. It has been noted that we see the largest gaps in comparative performance at the level of abstract relations. This observation is well grounded in the comparative data sampled here, but not just there. Charles Spearman, a major figure in the history of human intelligence testing, defined general intelligence as the ability to "educe relations" (Spearman, 1927), an insight into human intellectual variation that has been massively confirmed by decades of testing.

It may not be surprising that the means for dealing with just the transient contingencies of reinforcement has had lower evolutionary priority than the means for the enduring ones represented by concepts or by open-ended categories. But once an evolving species has a foothold at the level of abstract relations, the possibilities are unbounded. What then lies ahead are systematic descriptions, not just of transient contingencies of reinforcement, but of conditional ones. Language, logic, and mathematics, are the forms of abstract relations; knowledge itself is the substance.

References

- Barlow, H.B. (1972). Single units and sensation: A neuron doctrine for perceptual psychology? *Perception*, *1*, 371–394.
- Carter, D.E., & Werner, T.J. (1978). Complex learning and information processing by pigeons: A critical analysis. *Journal of the Experimental Analysis of Behavior*, *29*, 565–601.
- Commons, M.L., & Hallinan, P.W., with Fong, W., & McCarthy, K. (1990). Intelligent pattern recognition: Hierarchical organization of concepts. In M.L. Commons, R.J. Herrnstein, S.M. Kosslyn, & D.B. Mumford (Eds.), *Quantitative analyses of behavior: Computational and clinical approaches to pattern recognition and concept formation* (Vol. IX, pp. 127–153). Hillsdale, NJ: Erlbaum.
- 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.
- Delius, J.D., & Nowak, B. (1982). Visual symmetry recognition by pigeons. *Psychological Research*, *44*, 199–212.
- Dore, F.Y., & Dumas, C. (1987). Psychology of animal cognition: Piagetian studies. *Psychological Bulletin*, *102*, 219–233.
- Estes, W.K. (1986). Array models for category learning. *Cognitive Psychology*, *18*, 500–549.
- Gray, E.M. (1987). *Visual recognition of landmarks by the homing pigeon*. Unpublished undergraduate honors' thesis, Harvard University.
- Harlow, H.F. (1949). The formation of learning sets. *Psychological Review*, *56*, 51–65.
- Herrnstein, R.J. (1970). On the law of effect. *Journal of the Experimental Analysis of Behavior*, *13*, 243–266.
- Herrnstein, R.J. (1977). The evolution of behaviorism. *American Psychologist*, *32*, 593–603.
- Herrnstein, R.J. (1979). Acquisition, generalization, and discrimination reversal of a natural concept. *Journal of Experimental Psychology: Animal Behavior Processes*, *5*, 118–129.
- Herrnstein, R.J. (1982). Stimuli and the texture of experience. *Neuroscience & Biobehavioral Reviews*, *6*, 105–117.
- Herrnstein, R.J. (1984). Objects, categories, and discriminative stimuli. H.L. Roitblat, T.G. Bever, & H.S. Terrace (Eds.), *Animal cognition* (pp. 233–261). Hillsdale, NJ: Erlbaum.
- Herrnstein, R.J. (1985). Riddles of natural categorization. *Philosophical Transactions of the Royal Academy, London, B* *308*, 129–144.
- Herrnstein, R.J., & de Villiers, P.A. (1980). Fish as a natural category for people and pigeons. In G.H. Bower (Ed.), *The psychology of learning and motivation* (Vol. 14, pp. 59–95). New York: Academic Press.
- Herrnstein, R.J., Loveland, D.H., & Cable, C. (1976). Natural concepts in pigeons. *Journal of Experimental Psychology: Animal Behavior Processes*, *2*, 285–311.
- Herrnstein, R.J., Vaughan, W., Jr., Mumford, D.B., & Kosslyn, S.M. (1989). Teaching pigeons an abstract relational rule: Insideness. *Perception & Psychophysics*, *46*, 56–64.
- Hollard, V.D., & Delius, J.D. (1982). Rotational invariance in visual pattern recognition by pigeons and humans. *Science*, *218*, 804–806.
- Holmes, P.W. (1979). Transfer of matching performance in pigeons. *Journal of the Experimental Analysis of Behavior*, *31*, 103–114.
- Homa, D. (1984). On the nature of categories. In G.H. Bower (Ed.), *The psychology of learning and motivation* (Vol. 18, pp. 49–94). New York: Academic Press.
- Hubel, D.H., & Wiesel, T.N. (1962). Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *Journal of Physiology*, *160*, 106–154.
- Kamil, A.C. (1978). Systematic foraging by a nectar-feeding bird, the Amakihi (*Loxops virens*). *Journal of Comparative and Physiological Psychology*, *92*, 388–396.
- 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.

- Lombardi, C.M., Fachinelli, C.G., & Delius, J.D. (1984). Oddity of visual patterns conceptualized by pigeons. *Animal Learning and Behavior*, *12*, 2–6.
- Matsuzawa, T. (1985a). Colour naming and classification in a chimpanzee (*Pan troglodytes*). *Journal of Human Evolution*, *14*, 283–291.
- Matsuzawa, T. (1985b). Use of numbers by a chimpanzee. *Nature*, *315*, 57–59.
- Meyer, D.R., & Harlow, H.F. (1949). The development of transfer of response to patterning by monkeys. *Journal of Comparative and Physiological Psychology*, *42*, 454–462.
- Nickerson, R.S. (1965). Short-term memory for complex meaningful visual configurations: A demonstration of capacity. *Canadian Journal of Psychology*, *19*, 155–160.
- Osgood, C.E. (1953). *Method and theory in experimental psychology*. New York: Oxford University Press.
- Pepperberg, I.M. (1981). Functional vocalizations of an African grey parrot (*Psittacus erithacus*). *Zeitschrift für Tierpsychologie*, *55*, 139–151.
- Pepperberg, I.M. (1987). Evidence for conceptual quantitative abilities in the African parrot: Labeling of cardinal sets. *Ethology*, *75*, 37–61.
- Pfungst, O. (1911, reprinted in 1965). *Clever Hans (the horse of Mr. von Osten)* (C.L. Rahn, Trans.). New York: Holt, Rinehart and Winston.
- Pietrewicz, A.T., & Kamil, A.C. (1977). Visual detection of cryptic prey by blue jays (*Cyanocitta cristata*). *Science*, *195*, 580–582.
- Pisacreta, R., Lefave, P., Lesneski, T., & Potter, C. (1985). Transfer of oddity learning in the pigeon. *Animal Learning and Behavior*, *13*, 403–414.
- Pisacreta, R., Redwood, E., & Witt, K. (1984). Transfer of matching-to-figure samples in the pigeon. *Journal of the Experimental Analysis of Behavior*, *42*, 223–237.
- Premack, D. (1976). *Intelligence in ape and man*. Hillsdale, NJ: Erlbaum.
- Premack, D. (1986). *Gavagai! Or the future history of the animal language controversy*. Cambridge, MA: MIT Press.
- Real, P.G., Iannazzi, R., Kamil, A.C., & Heinrich, B. (1984). Discrimination and generalization of leaf damage by blue jays (*Cyanocitta cristata*). *Animal Learning and Behavior*, *12*, 202–208.
- Restle, F. (1959). A metric and an ordering on sets. *Psychometrika*, *24*, 207–220.
- Ristau, C.A., & Robbins, D. (1982). Language in the great apes: A critical review. *Advances in the Study of Behavior*, *12*, 141–225.
- Rosch, E.H. (1973). Natural categories. *Cognitive Psychology*, *4*, 328–350.
- Rumbaugh, D.M., & Pate, J.L. (1984). The evolution of cognition in primates: A comparative perspective. In H.L. Roitblat, T.G. Bever, & H.S. Terrace (Eds.), *Animal cognition* (pp. 569–587). Hillsdale, NJ: Erlbaum.
- Rumbaugh, D.M., Savage-Rumbaugh, S., & Hegel, M.T. (1987). Summation in the chimpanzee (*Pan troglodytes*). *Journal of Experimental Psychology: Animal Behavior Processes*, *13*, 107–115.
- Shepard, R.N. (1967). Recognition memory for words, sentences, and pictures. *Journal of Verbal Learning and Verbal Behavior*, *6*, 156–163.
- Shettleworth, S.J. (1972). Constraints on learning. In D.S. Lehrman, R.A. Hinde, & E. Shaw (Eds.), *Advances in the study of behavior* (Vol. 4, pp. 1–68). New York: Academic Press.
- Shettleworth, S.J., & Krebs, J.R. (1986). Stored and encountered seeds: A comparison of two spatial memory tasks in marsh tits and chickadees. *Journal of Experimental Psychology: Animal Behavior Processes*, *12*, 248–257.
- Smith, E.E., & Medin, D.L. (1981). *Categories and concepts*. Cambridge, MA: Harvard University Press.
- Spearman, C. (1927). *The abilities of man: Their nature and measurement*. New York: Macmillan.
- Terrace, H.S. (1979). *Nim*. New York: Knopf.
- Terrace, H.S. (1985). In the beginning was the “name.” *American Psychologist*, *40*, 1011–1028.
- Turney, T.H. (1982). The association of visual concepts and imitative vocalizations in the mynah (*Gracula religiosa*). *Bulletin of the Psychonomic Society*, *19*, 59–62.

- Tversky, A. (1977). Features of similarity. *Psychological Review*, *84*, 327–352.
- Ullman, S. (1984). Visual routines. *Cognition*, *18*, 97–159.
- Underwood, B.J. (1966). *Experimental psychology* (2nd ed.). New York: Meredith.
- Urcuioli, P.J. (1985). On the role of differential sample behaviors in matching-to-sample. *Journal of Experimental Psychology: Animal Behavior Processes*, *11*, 502–519.
- Vander Wall, S.B. (1982). An experimental analysis of cache recovery in Clark's nutcracker. *Animal Behavior*, *30*, 84–94.
- Vander Wall, S.B., & Balda, R.P. (1977). Coadaptations of the Clark's nutcracker and the pinon pine for efficient seed harvest and dispersal. *Ecology Monographs*, *47*, 89–111.
- Vaughan, W., Jr. (1988). Formation of equivalence sets in pigeons. *Journal of Experimental Psychology: Animal Behavior Processes*, *14*, 36–42.
- Vaughan, W., Jr., & Greene, S.L. (1983). Acquisition of absolute discriminations in pigeons. In M.L. Commons, R.J. Herrnstein, & A.R. Wagner (Eds.), *Quantitative analyses of behavior: Discrimination processes* (Vol. IV, pp. 231–238). Cambridge, MA: Ballinger.
- Vaughan, W., Jr., & Greene, S.L. (1984). Pigeon visual memory capacity. *Journal of Experimental Psychology: Animal Behavior Processes*, *10*, 256–271.
- Vaughan, W., Jr., & Herrnstein, R.J. (1987). Choosing among natural stimuli. *Journal of the Experimental Analysis of Behavior*, *47*, 5–16.
- von Uexküll, J. (1934). *Streifzüge durch die Umwelten von Tieren und Menschen*. Berlin: Springer. Translated by Claire H. Schiller as "A stroll through the worlds of animals and men", in C.H. Schiller (ed.) (1957). *Instinctive behavior: The development of a modern concept* (pp. 5–80). New York: International Universities Press.
- Wright, A.A., Santiago, H.C., Urcuioli, P.J., & Sands, S.F. (1983). Monkey and pigeon acquisition of same/different concept using pictorial stimuli. In M.L. Commons, R.J. Herrnstein, & A.R. Wagner (Eds.), *Quantitative analyses of behavior: Discrimination processes* (Vol. IV, pp. 295–317). Cambridge, MA: Ballinger.
- Younger, B.A., & Cohen, L.B. (1985). How infants form categories. In G.H. Bower (Ed.), *The psychology of learning and motivation* (Vol. 19, pp. 211–247). New York: Academic Press.
- Zentall, T.R., & Hogan, D.E. (1978). Same/different concept learning in the pigeon: The effect of negative instances and prior adaptation to transfer stimuli. *Journal of the Experimental Analysis of Behavior*, *30*, 177–186.