JOURNAL OF THE EXPERIMENTAL ANALYSIS OF BEHAVIOR 2002, **78,** 237–248 NUMBER 3 (NOVEMBER)

CATEGORIZATION, CONCEPT LEARNING, AND BEHAVIOR ANALYSIS: AN INTRODUCTION

THOMAS R. ZENTALL, MARK GALIZIO, AND THOMAS S. CRITCHFIELD

UNIVERSITY OF KENTUCKY,
UNIVERSITY OF NORTH CAROLINA AT WILMINGTON,
AND ILLINOIS STATE UNIVERSITY

Categorization and concept learning encompass some of the most important aspects of behavior, but historically they have not been central topics in the experimental analysis of behavior. To introduce this special issue of the *Journal of the Experimental Analysis of Behavior* (*JEAB*), we define key terms; distinguish between the study of concepts and the study of concept learning; describe three types of concept learning characterized by the stimulus classes they yield; and briefly identify several other themes (e.g., quantitative modeling and ties to language) that appear in the literature. As the special issue demonstrates, a surprising amount and diversity of work is being conducted that either represents a behavior-analytic perspective or can inform or constructively challenge this perspective.

Key words: categorization, concept learning, stimulus class, function transfer

> Categorization is not a matter to be taken lightly. There is nothing more basic than categorization to our thought, perception, action, and speech. (Lakoff, 1987, p. 5)

> Concepts give our world stability. They capture the notion that many objects or events are alike in some important respects, and hence can be thought about and responded to in ways already mastered. Concepts also allow us to go beyond the information given; for once we have assigned an entity to a class . . . we can infer some of its . . . attributes. (Smith & Medin, p. 1)

There is, perhaps, no larger or more diverse literature within experimental psychology than that focused on categorization and concept learning. This topic is, to the casual observer, most directly associated with human cognitive psychology, within which the largest volume of research and theory building has taken place. Cognitive psychologists Laurence and Margolis (1999), in a book reviewed in the present issue, minced no words about the association: ''Concepts are the most fundamental constructs in theories of the mind'' (p. 3). Yet scholars from many research communities have struggled to come to grips with the complex repertoires that the topic encompasses.

The heterogeneity of this research area is evident in the absence of a consensus definition of the term *concept* (see Palmer, this issue; Wasserman & Bhatt, 1992). Writers tend to stress the importance of concepts rather than specifying their defining features—perhaps, as Palmer speculates in his review of Margolis and Laurence (1999), ''regarding the term as too familiar to need definition'' (p. 598). Nevertheless, an introduction to this special issue demands at least an attempt to define its subject matter.

Typically in cognitive psychology, *categorization* is regarded as a process of determining what things ''belong together,'' and a *category* is a group or class of stimuli or events that so cohere. A *concept* is thought to be knowledge that facilitates the categorization process (e.g., Barsalou, 1991, 1992). Consistent with the representational style of much cognitive theorizing, conceptual knowledge is often portrayed as existing independently of any particular behavior–environment relation. This is assumed partly because, once a categorization repertoire is in place, an individual may be able to categorize both previously en-

---

Order of authorship for this article was determined by random drawing. We are grateful to the numerous reviewers, who met the dual challenge of evaluating manuscripts according to *JEAB*'s usual expectations while embracing the conceptual and methodological diversity inherent in the topic, and to the authors of the articles contained herein, many of whom invested extra effort in tailoring their work to an unfamiliar audience.

Address correspondence to Tom Zentall at the Department of Psychology, University of Kentucky, Lexington, Kentucky 40506 (e-mail: zentall@pop.uky.edu); Mark Galizio at the Department of Psychology, University of North Carolina at Wilmington, Wilmington, North Carolina 28403 (e-mail: galizio@uncwil.edu); or Tom Critchfield at the Department of Psychology, Illinois State University, Normal, Illinois 61704 (e-mail: tscritc@ilstu.edu).

countered stimuli and novel events, suggesting to some observers that the latter are recognized via comparison to general information represented in memory. Thus, the goal of many studies in cognitive psychology is to map the knowledge that humans presumably apply in already established patterns of categorization. For example, structured interviewing and other techniques may be used to determine what entities people include in a category like *birds*; which of these entities are considered to be more or less typical of the category; and whether hierarchical relations apply to the category or instances within it (e.g., Rosch, 1978).

Behavior analysts are likely to regard as foreign this practice of describing terminal performance without examining the necessary and sufficient conditions for its emergence. Reservations about the approach are warranted. Although cognitive psychologists expend much energy debating the structure and contents of the knowledge that is assumed to underpin categorization and the means by which it is compared to new perceptual experiences (Laurence & Margolis, 1999; see Palmer, this issue, for a brief synopsis of some relevant theories), their accounts can be difficult to distinguish empirically. Barsalou (1992) has noted a tendency for competing cognitive theories to make similar predictions and to account equally well for data obtained from human subjects. Perhaps more important for present purposes, this focus on knowledge may discourage attention to the role of experience in creating and maintaining conceptual behavior (Astley & Wasserman, 1996).

AN OPERATIONAL APPROACH TO CONCEPT LEARNING

From a behavior-analytic perspective, the present topic provides an opportunity to apply the operational analysis of psychological terms that Skinner (e.g., 1945) frequently espoused. Rather than speculating about the status of hypothetical knowledge structures, it is possible to examine the circumstances under which we speak of conceptualization—that is, what individuals are doing when they are said to behave conceptually, and how they came to behave in that way. Keller and Schoenfeld (1950) specified this very point of departure by suggesting that ''when a group of objects gets the same response, when they form a class the members of which are reacted to similarly, we speak of a concept'' (p. 154). Thus, categorization may be said to incorporate a pattern of systematic differential responding to classes of nonidentical, though potentially discriminable, stimuli (see Fields, Reeve, et al., this issue; Wasserman & Bhatt, 1992). A category is a class of stimuli that occasion common responses in a given context. Such classes include stimuli involved in an explicit learning history plus, potentially, novel stimuli to which the fruits of this history may transfer. Many writers use the terms *category* and *stimulus class* more or less interchangeably; we will follow that practice here.

When the stimuli within and between categories vary along relatively simple dimensions (e.g., wavelength, size, brightness), categorization is readily conceived in the same terms as stimulus discrimination and generalization. For example, ''Generalization *within* classes and discrimination *between* classes—this is the essence of concepts'' (Keller & Schoenfeld, 1950, p. 155). The analytical challenge becomes more daunting, of course, as category membership is determined more complexly (e.g., Herrnstein, 1990). Consider, as an instructive case, the *balan* category of the Australian aboriginal language Dyirbal, which ''includes women, fire, and dangerous things. It also includes birds that are *not* dangerous, as well as exceptional animals such as the platypus, bandicoot, and echidna'' (Lakoff, 1987, p. 5). Any plausible account must also explain how categories add and lose members, merge and fracture, share members that may belong to different categories under different circumstances, support the spontaneous transfer of function from one member to another, and so forth.

Some writers have gone so far as to label the capacity to glean abstract relations, such as those that apparently unite many categories, as the essence of what it means to be human (e.g., see Deacon, 1997). Regardless of whether conceptual repertoires are uniquely human (many articles in the present issue suggest that they are not), they are clearly among the most interesting behavioral phenomena available for study. Because of their richness, generativity, and adaptability, they invite a thorough experimental and theoretical analysis.

Unfortunately, categories and concepts have been addressed only sporadically within behavior analysis. This neglect may be understood partly as a rejection of the cognitive theoretical worldview that has helped to define the topic (in the present issue, Palmer catalogues some key points of contention between behavior-analytic and cognitive views). But taking issue with the theoretical perspective that has dominated the concepts and categories literature does not render the relevant behavioral phenomena any less provocative. Moreover, objection through silence persuades no one. Although Keller and Schoenfeld (1950) provided the outlines of a behavioral analysis of conceptual behavior more than half a century ago, scholars outside behavior analysis have scarcely noticed, and they are unlikely to do so in the absence of persuasive empirical evidence (Schwartz, Wasserman, & Robbins, 2002).

The needed research requires a clearly defined subject matter, and we have already suggested that concepts, defined largely in terms of abstract knowledge, provide a slippery foundation for an experimental analysis. Fortunately, ‘‘knowledge’’ can be regarded, not as an entity, but as a linguistic surrogate for learning histories that can be operationalized and, in many cases, studied experimentally (e.g., see Gagné, this issue; Maddox, this issue; Skinner, 1977; Wixted & Gaitan, in press). In search of more secure theoretical footing, the authors of most articles in the present issue have, in some fashion, recast the notion of knowledge in terms of necessary and sufficient conditions for conceptual behavior. This subtle shift in emphasis, from concepts to *concept learning*, diverts attention at least partly away from taxonomizing knowledge and toward identifying the functional relations between behavior and environment that provide the basis for conceptual behavior.

The preceding insights are not, of course, unique to behavior analysis. Many research traditions (including human cognitive psychology; see Maddox, this issue) include attempts to create artificial categories uniting experimenter-selected rather than everyday stimuli. Almost universally, these efforts require subjects to respond, with feedback, to examples and nonexamples of category members (Barsalou, 1991). The goal is ‘‘to establish high degrees of control over category knowledge’’ (Barsalou, 1992, p. 31), with knowledge operationalized more or less as above. Such paradigms, which are comfortingly reminiscent of discrimination learning procedures, form the basis of most of the research described in the present issue and constitute an important point of contact between diverse communities of researchers.

Although the scientific community with which *JEAB* is most associated has not made categorization and concept learning the centerpiece of its empirical and theoretical contributions, a remarkable amount of work is being conducted that expressly represents a behavior-analytic perspective, can inform this perspective, or can constructively challenge this perspective. The overarching purpose of this special issue, therefore, is to bring together the best current work of behavioral researchers who seek to determine which species are capable of which conceptual repertoires, and what experiences and circumstances make these repertoires possible.

We anticipate that some *JEAB* readers will have only passing acquaintance with this research area. For these readers, we seek to provide a cross section of contemporary research efforts and illustrate something of the breadth and value of this area of study. To place this work into a broad context, below we provide some further orienting remarks about the area. Because the literature on categorization and concept learning is so vast, even those who work in this area may be unfamiliar with the efforts of other investigators. A key goal of this special issue is to promote a cross propagation of empirical and theoretical insights, and to demonstrate that *JEAB* can be the locus for a productive exchange on a rich and compelling topic.

TYPES OF CONCEPTS

The work described in the present issue draws from multiple scholarly communities. One community of investigators examines categorization and concept learning in nonhumans, both to plumb the abilities of specific species and to assess interspecies similarities and differences. A second community of researchers has been concerned with the formation of stimulus classes (including equivalence classes) primarily in humans. Much of their work bears on questions about categorization and concept learning, and some of it addresses these topics explicitly. Finally, a large community of cognitive psychologists seeks to understand categories and concepts in humans. To avoid unnecessary entanglement in theoretical debates between and within communities, below we emphasize three broad types of relations that appear to unite events within a category. In *perceptual concepts*, stimuli are grouped primarily on the basis of shared physical features. In *relational concepts*, it is not the physical features of stimuli per se but the relations among these features that are grouped. Finally, in *associative concepts*, stimuli are grouped on the basis of shared function (e.g., a common response that they engender, or a common consequence with which they are correlated).

Perceptual Concepts

The classes of stimuli that are united in perceptual concepts may be said, from a subject's perspective, to bear physical similarity to one another. Because stimuli often are said to be similar on the circular basis that subjects respond similarly to them, similarity is construed here as relatively little separation along a well-defined physical dimension. Straightforward examples can be found in the present issue in articles by Maddox (line length) and by Fields and colleagues (gradations between two photographs as altered by commercial morphing software). Although most laboratory concept tasks involve static stimuli, Herbranson, Fremouw, and Shimp (this issue) show that pigeons can accurately categorize dynamic properties such as stimulus movement and direction.

In the interest of modeling categorization as it occurs in the natural environment, many studies employ stimuli that are not defined by a few simple features. For example, research with nonhumans involves the discrimination of photographs or drawings that contain a particular type of object, such as a person, from those that do not contain the object. Herrnstein and Loveland (1964) taught pigeons to discriminate photographs containing a person from photographs that did not (in such a case, category membership may be said to be determined, in the experimental contingencies, by what the experimenter recognizes as an image of a person). Following this kind of training, subjects may show a high degree of class-consistent responding to novel exemplars of the stimulus sets (Herrnstein, Loveland, & Cable, 1976). In the present issue, Vonk and MacDonald extend this type of procedure to gorillas, whose conceptual abilities have been studied infrequently.

In studies like that of Herrnstein and Loveland (1964), the pictures in the positive set (to which responding is reinforced) and in the negative set (to which responding is extinguished) vary in terms of perspective, background, color, number of relevant items portrayed, and so forth. In such cases, as Vonk and MacDonald (this issue) note, it is more difficult to think in terms of simple stimulus control processes. One noteworthy feature of the Vonk and MacDonald study is an extensive, and apparently unsuccessful, attempt to identify simple stimulus features that guide categorization. In such cases, stimulus variation within a picture set may be as great as the variation between the sets and, when many pictures are used (Herrnstein & Loveland, 1964, used several hundred), it seems unlikely that each separate picture becomes an independent discriminative stimulus.

Even when complex stimuli are employed, however, concept learning mirrors discrimination learning in tantalizing ways. For example, in the acquisition of simple discriminations, experience with both positive and negative stimuli appears to be necessary; establishing a response in the presence of one stimulus does not ensure differential responding when other stimuli are introduced (e.g., Newman & Baron, 1965). Perhaps analogously, training animals to respond to a single *set* of stimuli may be insufficient to ensure differential responding among multiple sets. Sutton and Roberts, in the present issue, report that discrimination training between categories is needed for pigeons to respond differentially to untrained members of the training set versus untrained members of a different training set (i.e., category distinctions may not be made until members of at least two categories are compared). Thus, experience with both ''examples'' and ''nonexamples'' appears to be essential.

Relational Concepts

Compared to perceptual concept learning, in which absolute properties of the stimuli may guide responding, relational concept learning makes use of more abstract properties of the stimuli. One of the simplest and most studied relational concepts is same versus different. Same–different concept learning can be studied using conditional discriminations in which subjects must respond to the comparison that matches the sample. For example, after pigeons have been trained to match two shapes, they show facilitated acquisition, relative to an appropriate control, when trained with two novel hues (Zentall & Hogan, 1978). Even better transfer can be found when pigeons are trained on an identity task with a large number of stimuli (Wright, Cook, Rivera, Sands, & Delius, 1988).

Perhaps the strongest evidence of same–different learning by pigeons comes from research in which the simultaneous display of similar objects is discriminated from displays in which all of the objects are clearly different from each other. Pigeons trained in this way learn to discriminate the training exemplars and also to discriminate novel stimuli involving the same relations as in training (e.g., Cook, Katz, & Cavoto, 1997). In the present issue, Cook reports that the resulting same–different categories are general and can include stimuli that vary along different perceptual dimensions such as texture, feature, geometric shape, naturalistic drawing, or an object depicted in a photograph. Also in the present issue, Wasserman, Young, and Peissig report that detecting same and different arrays occurs very rapidly and thus does not appear to require the sequential comparison of the stimuli in the array.

There is also evidence that some species are capable of second-order same–different learning, in which the *relation* between two objects must be matched rather than the objects themselves. With this procedure, the sample consists of a pair of objects that are either identical or different, and the correct comparison consists of two stimuli that are different from the sample but that bear the same relation to each other as those in the sample. For example, if shown Sample AA and given a choice between BB and CD, the reinforced choice would be BB. Thus, it is not the match between stimuli but the match between the stimulus–stimulus relations that must be learned. Chimpanzees acquire such an abstract conditional discrimination and can transfer it to novel exemplars (Thompson, Oden, & Boysen, 1997). In the present issue, possibly analogous performances are shown in human ‘‘equivalence–equivalence responding’’ (Stewart, Barnes-Holmes, Roche, & Smeets).

Associative Concepts

In associative concept learning, the stimuli within classes bear no obvious physical similarity to one another, but rather cohere because of shared functional properties. Sidman’s (e.g., 1994) seminal work on stimulus equivalence has generated enormous interest, in the behavior-analytic community, in the study of non-similarity-based stimulus classes. This interest arose, in part, because arbitrary relations were thought to provide a tool for studying symbolic processes relevant to language and cognition (e.g., Sidman, 1971, 2000). As research on stimulus equivalence has evolved into a major focus in the experimental analysis of human behavior, researchers of animal learning and cognition have been investigating similar phenomena in nonhumans.

A brief review of the basic methods used to study stimulus equivalence, as defined by Sidman (e.g., 1994), will illustrate one way that associative concept learning can be studied in the laboratory. Typically, arbitrary match-to-sample training is used to establish at least two conditional discriminations involving physically unrelated stimuli. On each trial, the correct comparison is conditional on the particular sample presented. For example, given Sample Stimulus A1, a choice of Comparison Stimulus B1 is reinforced, and given Sample A2, choosing B2 is reinforced. Another set of relations could then reinforce a choice of C1 given Sample B1 and C2 given Sample B2. After such training, three kinds of untrained performances often emerge. Sidman designated these reflexivity (e.g., given A1 as the sample, the subject chooses Comparison A1), symmetry (e.g., given B1 as the sample, the subject chooses Comparison A1), and transitivity (e.g., given A1 as the sample, the subject chooses Comparison C1).

Humans have shown the emergence of reflexivity, symmetry, and transitivity without explicit reinforcement of these relations, and, historically, the demonstration of all three relations was considered to be the definition of an equivalence class (e.g., Sidman, 1994; Sidman & Tailby, 1982). Over the past 25 years or so, many studies have been directed toward evaluating the mechanics of equivalence class formation (e.g., types of training and testing formats, preconditions of class merger, expansion, and fracture; and the stimulus control processes that promote or preclude class formation; see Arntzen & Holth, 1997; Carrigan & Sidman, 1992; Lane & Critchfield, 1998; Pilgrim & Galizio, 1995; Saunders & Green, 1999; Sidman, Kirk, & Willson-Morris, 1985).

Some commentators have wondered whether stimulus equivalence is a uniquely human phenomenon (e.g., see Hayes, 1989), because research has suggested that not all of the formal properties of equivalence classes (as defined by Sidman & Tailby, 1982) are readily demonstrated in nonhumans. For example, symmetrical relations may fail to emerge in nonhumans after training involving arbitrary stimuli (Sidman et al., 1982). One possible explanation of this failure emphasizes not a generic deficit but rather competing control by factors such as stimulus location. In the present issue, Lionello-DeNolf and Urcuioli describe an experiment in which pigeons were tested for symmetry after multiple sample-location training designed to reduce control by a particular location. Symmetrical responding was directly reinforced with some stimulus sets before testing for symmetry with new sets. Although baseline matching transferred to novel stimulus locations, no evidence of the emergence of symmetry was obtained in any of their experiments.

As the preceding case illustrates, recent years have seen a convergence of interests between equivalence researchers and researchers who study associative concepts in animal cognition (Zentall & Smeets, 1996). Some studies influenced by this convergence now suggest the capacity of certain nonhuman species to acquire equivalence relations (see Kastak & Schusterman, this issue; Kastak, Schusterman, & Kastak, 2001). Many more studies have shown functional substitutability of class members, leaving the distinctions between equivalence and functional classes blurred at best (see Sidman, 2000). To illustrate, we consider some of the procedures through which such functional classes have been examined.

Common-response training. Stimuli can become united into a class by virtue of association with a common response. Many-to-one conditional discrimination procedures, for example, arrange the requisite common response by reinforcing the matching of two or more samples (e.g., a red field and vertical lines) to the same comparison selection (e.g., a large circle), and different samples (e.g., a green field and horizontal lines) to another comparison selection (e.g., a small circle). Several lines of evidence indicate that in nonhumans a relation develops between the samples associated with a common comparison (e.g., Zentall, Sherburne, & Urcuioli, 1993; Zentall, Steirn, Sherburne, & Urcuioli, 1991), just as is the case for humans (e.g., Arntzen & Holth, 1997; Fields, Reeve, et al., this issue; see also McDaniel, Nuefeld, & Damico-Nettleton, 2001).

A variety of means exist to assess the associative relations that emerge in nonhumans following many-to-one training. For example, if Samples A and B share a common comparison selection, C, and further training relates Sample A to new Comparison D, Sample B will be matched to Comparison D without additional training (Urcuioli, Zentall, Jackson-Smith, & Steirn, 1989). Another approach involves comparing the discriminability of samples associated with a common comparison selection with that of samples associated with different comparison selections. If relations have formed between samples associated with the same comparison selection, then it should be harder to discriminate between those samples than between samples associated with different comparison selections. Recent research suggests that this is indeed the case (Kaiser, Sherburne, Steirn, & Zentall, 1997).

Common-sample training. It is worth noting that, in humans, associative concepts develop fairly readily through one-to-many training, in which a common sample is associated with two or more comparison selections (e.g., Arntzen & Holth, 1997). Surprisingly, similar results have not been found with pigeons, even when conditions are otherwise similar to those under which associative concepts have been shown to develop during many-to-one training (Urcuioli & Zentall, 1993). The asymmetry of these effects hints at differences in the way humans and other animals form these associative concepts, although interspecies similarities may yet exist. In the present issue, Fields, Reeve, et al. report that the likelihood of humans acquiring a *generalized* catergorization repertoire (one that applies across several different mixed perceptual-associative categories) is positively related to the number of sample stimuli used during training and is negatively related to the number of comparison stimuli. Their findings suggest that one-to-many training may interfere with generalized categorization in humans.

Common-outcome training. Stimuli can cohere into classes by way of their association with a common outcome. For example, Edwards, Jagielo, Zentall, and Hogan (1982) trained pigeons on two identity-matching tasks, one with lines and the other with hues. In each task, correct choices of one of the comparisons (e.g., red and vertical) were reinforced with peas, whereas correct choices of the other comparisons (e.g., green and horizontal) were reinforced with wheat. Following training, pigeons showed positive transfer to trials in which hues and lines associated with a common outcome could be matched. For analogous findings with humans, see Dube, McIlvane, Maguire, Mackay, and Stoddard (1989).

Serial reversals. Another procedure for developing stimulus classes is to treat a set of stimuli the same through a number of transformations (Vaughan, 1988). Vaughan randomly assigned photographs of trees to two arbitrary sets, A and B, and then trained pigeons to respond to those in Set A but not to respond to those in Set B. Following training, the valence associated with each set was reversed, and then reversed again, repeatedly. Across reversals of stimulus sets, the number of trials needed for responding to follow suit decreased to only a few trials, suggesting that these arbitrarily assigned stimuli had become two functional stimulus classes, in spite of the fact that stimuli differed as much within sets as between sets in terms of physical similarity and reinforcement histories. In the present issue, Jitsumori, Siemann, Lehr, and Delius report that similar results can be obtained using simultaneous discriminations. The use of simultaneous discriminations permits the presentation of novel pairings of the training stimuli. For example, the positive stimulus from one training pair could be presented with the negative stimulus from a different training pair.

Symmetry training. Recent research suggests that symmetry training, in which A → B and B → A relations are taught, can serve to form a stimulus class consisting of those stimuli. Zentall, Clement, and Weaver (in press) found that, following this training, if one stimulus was also associated with a new stimulus (e.g., B → C) in a conditional discrimination, the untrained relation A → C emerged as well.

Choice by exclusion. When animals learn to match, it is assumed that they learn to select the comparison based on its acquired association with the sample. Articles in the present issue indicate that if chimpanzees or sea lions are presented with a novel sample and one novel and one familiar comparison (i.e., a comparison already associated with a different sample), the animals will tend to choose by exclusion (Beran & Washburn; Kastak & Schusterman). That is, they will avoid the familiar comparison (that with an experimentally defined function) and choose the novel (undefined) one. Control procedures reveal that this is not simply a preference for novel comparisons but rather a general capacity to match undefined stimuli. Such choices therefore suggest an indirect measure of concept learning.

Transfer of function. In many of the preceding examples, concept learning was evaluated through selection-based repertoires involving the matching of one class member to another. Such repertoires have been described as part of a more general transfer of function (also called transformation of function or inheritance of meaning), in which the stimuli within a class spontaneously share whatever functions each has acquired separately (Dougher & Markham, 1996; Dymond & Rehfeldt, 2000). Manabe, Kawashima, and Staddon (1995) described a procedure in which, to produce comparison stimuli during match-to-sample training, budgerigars were required to make distinct vocal responses to each of two samples. Later, new samples were associated with the same comparisons (creating many-to-one training), but without differential vocalization requirements. Differential vocalizations spontaneously transferred to the new stimuli, apparently on the basis of association with common comparison selections. In the present issue, Urcuioli et al. describe a successful attempt to replicate this effect in pigeons, substituting high-rate versus low-rate pecking patterns for differential vocalization.

Transfer of function also has been a focus in many studies of stimulus equivalence. For example, Dougher and colleagues first created classes of arbitrary stimuli using stimulus equivalence procedures, then paired shock with one member of a class, making it a conditioned stimulus for skin conductance responses indicating fear elicitation. The elicitation function appeared spontaneously in other members of the same class (Dougher, Augustson, Markham, Greenway, & Wulfert, 1994). As Gagné (this issue) suggests, such transfer-of-function outcomes are reminiscent of the shared meaning evident in hybrid lexical concepts. Gagné proposes an experiential basis for this transfer of function to which readers may readily apply familiar behavior principles (e.g., conditional relations and the matching law).

In general, the circumstances that facilitate transfer of function are poorly understood (e.g., see Barnett & Ceci, 2002; Barsalou, 1992; Dymond & Rehfeldt, 2000). In the present issue, Urcuioli et al. offer an extended analysis of conditions under which such effects might arise, and address the possibility that adventitious reinforcement might create spurious cases of function transfer. Also in this issue, Markham and Markham explore some of the necessary and sufficient conditions for transfer of function in humans. They describe a case in which stimulus class formation (on the basis of common respondent functions) apparently failed, but an operant response function, trained to one stimulus, propagated through the putative class nevertheless. The results provide fodder for interesting speculation about what exactly creates and defines stimulus classes.

OTHER ISSUES

Theoretical Views

Much research on categorization and concept learning is guided by cognitive theories that run counter to traditions in behavior analysis (Palmer, this issue), but the cognitive literature merits attention for at least two reasons. First, regardless of its theoretical underpinnings, this literature describes many empirical phenomena that any noncognitive theory must explain (e.g., see *Concepts and Language,* below). Second, cognitive psychology is not monolithic, and behavior analysts may be surprised to find that they share many assumptions with some of their cognitive counterparts. In the present issue, Maddox and Gagné provide cases in point. Maddox, for instance, stresses the importance of providing individuals with extended exposure to experimental manipulations and of evaluating individual-subject response functions. His quantitative models of perceptual concept learning draw heavily on signal-detection principles and the effects of consequences on behavior.

Readers of research on stimulus relations (such as equivalence) in humans will discover the influence of three main theories. Sidman's (2000) stance is that equivalence relations represent a fundamental outcome of a reinforcement contingency such that all elements of the contingency (conditional and discriminative stimuli, responses, and reinforcers) are posited to become equivalence class members. Several papers in the present issue incorporate Sidman's framework (e.g., Fields et al.; Griffee & Dougher). Two competing views—naming theory (Horne & Lowe, 1996) and relational frame theory (Hayes, Barnes-Holmes, & Roche, 2001)—both propose that the emergence of equivalence relations reflects previously established higher order operant behavior.

In the case of naming theory, the prerequisite operant is a bidirectional relation between objects and the speaker-listener behavior (names) they occasion (note that naming, construed somewhat differently, also plays a role in some cognitive accounts; see Markman, 1991). An example of research conducted within the framework of naming theory is provided in present issue by Lowe, Horne, Harris, and Randle, who describe a variation of the common-response procedure in which children were trained to make one common naming response to a set of physically unrelated stimuli and another common naming response to a different set of stimuli. The common names were sufficient to establish stimulus classes determined by sorting the stimuli in a category match-to-sample test.

Relational frame theory posits that much complex behavior, including equivalence, is better understood as the result of ''arbitrarily applicable relational responding'' called a *relational frame* (relational frames also are invoked in some cognitive accounts; see Barsalou, 1993, and Gagné, this issue). From this vantage point, equivalence is one of many possible relational operants that Hayes et al. (2001) call a *frame of coordination.* They use the terms *mutual entailment* and *combinatorial entailment,* rather than *symmetry* and *transitivity,* because in some relational frames (e.g., ''greater than'') entailed relations are different from those that emerge from equivalence relations. Thus, relational frame theory is designed to account for a broad range of relations including equivalence phenomena. Some aspects of the theory are illustrated in the attempt by Stewart et al. (this issue) to craft a behavior-analytically based model of analogy.

It is worth noting that both relational frame theory and naming theory have been criticized for vagueness about the origins of higher order operants that supposedly underpin stimulus class formation. Although not targeting these theories per se, Fields, Reeve, et al. (this issue) describe some of the training conditions under which a generalized repertoire for categorizing perceptually defined stimuli is likely to arise. Their experiment illustrates the feasibility of conducting studies that bear on the plausibility of accounts based on higher order operants.

Formal Modeling

Some investigations of concept learning test, or support the development of, formal models. In human cognitive psychology, this is most evident in connectionist paradigms that focus on computer simulations of neural networks (e.g., Barsalou, 1993). Computer simulations also play a role in some areas of concept learning research with nonhumans (e.g., Wynne, 1995). Quantitative modeling, in the style typical of many investigations published in *JEAB,* has not played a large role in the study of conceptual behavior, but cases worthy of examination can be identified in human cognitive psychology. Gluck and Bower (1988), for example, presented a model of concept learning with roots in the Rescorla-Wagner (1972) account of respondent conditioning. In the present issue, Maddox describes a quantitative model of learning perceptual categories, and Gagné describes a model of conceptual combination.

Concepts and Language

Conceptual behavior often is thought to be linked to complex human capabilities such as language, and much research has examined arbitrary stimulus relations in humans as a theoretical basis of language and cognition (see Hayes et al., 2001; Horne & Lowe, 1996; Sidman, 1971, 2000). This work is in its infancy, however, and it remains to be seen how adequately laboratory-generated stimulus relations can simulate, and support accounts of, everyday language and cognition. It is encouraging that several phenomena typically observed in lexical classes have been modeled with equivalence relations. For example, demonstrations of category clustering in free recall (Galizio, Stewart, & Pilgrim, 2001), fast lexical mapping (Wilkinson, Dube, & McIlvane, 1996), and semantic priming (Hayes & Bisset, 1998) have all been accomplished with equivalence class procedures.

Several papers in the present issue extend such analyses to additional aspects of language-based categories. For example, lexical classes often include both members that are related perceptually and members that are not (e.g., Lane, Clow, Innis, & Critchfield, 1998). The category *furniture* includes exemplars such as chairs that are perceptually related to one another, but might not be perceptually related to other category members like cabinets or beds. Fields et al., in two papers in the present issue, explore the merger of perceptually based classes with equivalence classes in generalized equivalence classes.

Language-based categories often take on a hierarchical structure involving nonequivalence relations between members. For instance, all birds have wings but not all wings are found on birds; all primates are mammals, but not all mammals are primates; and so forth. In the present issue, Griffee and Dougher use differential reinforcement contingencies to establish hierarchical category relations that provide an elegant model of superordinate, subordinate, and basic level categorization. The Stewart et al. model of analogy in this issue provides another relevant example.

The preceding cases all involve attempts to generalize from concept learning, as arranged in the laboratory, to naturally occurring language phenomena outside the laboratory. In the present issue, Gagné illustrates the opposite approach, in which naturally occurring abilities are brought into the laboratory and experimentally analyzed. Gagné's research promotes testable predictions about the learning histories that may underlie conceptual combination involving the merger of lexical categories through noun–noun combinations (e.g., *exam headache*).

As Gagné's article indicates, members of lexical classes often exchange multiple functions, and transfer of function may not always reciprocate among stimuli within a class (see also Rosch, 1978). Unfortunately, data relevant to these complex effects are scarce in the literature on stimulus class formation, making an example from Clow (2000) worthy of mention. After humans had acquired equivalence classes, separate training procedures established additional discriminative functions for two class members. These additional functions were not mutually exclusive, potentially allowing the blending of functions through transfer. Class members to which no discriminative function had been explicitly trained almost always inherited one such function, but conjunctive transfer, reflecting the blending of discriminative repertoires, occurred in only about half of these cases. Stimuli to which one discriminative function had been explicitly trained almost never acquired a function that had been trained to another stimulus in the same class. These outcomes broadly mimic patterns seen in lexical classes, and they highlight an important possible contribution of stimulus class research. Studies of conceptual combination almost always focus on listener behavior (i.e., comprehension; see Gagné, this issue) or do not distinguish clearly between speaker behavior (production) and listener behavior (e.g., Barsalou, 1991). Function-transfer procedures in stimulus class research provide a possible means of modeling speaker effects.

CONCLUSION

We have presented a brief introduction to some of the problems that have concerned investigators of categorization and concept learning. We did not attempt a comprehensive survey of this research area because the relevant literature is too extensive. Consider, for example, that the Margolis and Laurence (1999) volume reviewed in this issue, which includes only selected seminal papers from only one research community, consists of over 600 densely packed pages. The present discussion is intended to place the articles of the special issue into a context of interest to journal readers, but it barely scratches the surface of a complex topic. One goal of the special issue, therefore, is to promote further reading in this area.

For a review of the literature on categorization and concept learning in nonhumans, we recommend Roberts (1998, chap. 11) and Honig and Fetterman (1992). For an introduction to the research on stimulus equivalence in humans, we recommend Sidman (1994). For discussion of the common research agendas that have united animal learning and human equivalence communities, we suggest a volume edited by Zentall and Smeets (1996; see also a 1993 special issue of *The Psychological Record,* Vol. 43, No. 4). Finally, no consideration of categorization and concept learning will be complete without some attention to the voluminous literature in human cognitive psychology. Barsalou (1992, chap. 2 and 7) provides a brief, accessible introduction to mainstream cognitive work. Extensive surveys can be found in Smith and Medin (1981) and the Margolis and Laurence (1999) volume reviewed in the present issue.

We hope that, ultimately, this special issue will encourage new applications of the experimental analysis of behavior to the phenomena and processes involved in categorization and concept learning. Systematic programs of research that complement, challenge, and extend those sampled in this special issue are called for, and the reports collected here illuminate some of the many questions worth addressing. If nothing else, the present issue should demonstrate that, no matter how vast the community of scholars who investigate categorization and concept learning, there is work enough for all who are interested.

REFERENCES

Arntzen, E., & Holth, P. (1997). Probability of stimulus equivalence as a function of training design. *The Psychological Record, 47,* 309–320.

Astley, S. L., & Wasserman, E. A. (1996). Mediating associations, essentialism, and non-similarity-based categorization. In T. R. Zentall & P. M. Smeets (Eds.), *Stimulus class formation in humans and animals* (pp. 111–133). Amsterdam: Elsevier.

Barnett, S. M., & Ceci, S. J. (2002). When and where do we apply what we learn? A taxonomy for far transfer. *Psychological Bulletin, 128,* 612–637.

Barsalou, L. W. (1991). Deriving categories to achieve goals. In G. H. Bower (Ed.), *The psychology of learning and motivation* (Vol. 27, pp. 1–64). San Diego: Academic Press.

Barsalou, L. W. (1992). *Cognitive psychology: An overview for cognitive scientists.* Hillsdale, NJ: Erlbaum.

Barsalou, L. W. (1993). Flexibility, structure, and linguistic vagary in concepts: Manifestations of a compositional system of perceptual symbols. In A. F. Collins, S. E. Gathercole, M. A. Conway, & P. E. Mathis (Eds.), *Theories of memory* (pp. 29–89). Hillsdale, NJ: Erlbaum.

Carrigan, P. F., Jr., & Sidman, M. (1992). Conditional discrimination and equivalence relations: A theoretical analysis of control by negative stimuli. *Journal of the Experimental Analysis of Behavior, 58,* 183–204.

Clow, J. K. (2000). The combination of response functions in merged equivalence classes: A behavioral paradigm for conceptual combination. *Dissertation Abstracts International, 61(3B),* 1619.

Cook, R. G., Katz, J. S., & Cavoto, B. R. (1997). Pigeon same-different concept learning with multiple stimulus classes. *Journal of Experimental Psychology: Animal Behavior Processes, 23,* 417–433.

Deacon, T. W. (1997). *The symbolic species: The co-evolution of language and the brain.* New York: Norton.

Dougher, M. J., Augustson, E. M., Markham, M. R., Greenway, D., & Wulfert, E. (1994). The transfer of respondent eliciting and extinction functions through stimulus equivalence classes. *Journal of the Experimental Analysis of Behavior, 50,* 125–144.

Dougher, M. J., & Markham, M. R. (1996). Stimulus classes and the untrained acquisition of function. In T. R. Zentall & P. M. Smeets (Eds.), *Stimulus class formation in humans and animals* (pp. 137–152). Amsterdam: Elsevier.

Dube, W. V., McIlvane, W. J., Maguire, R. W., Mackay, H. A., & Stoddard, L. T. (1989). Stimulus class formation and stimulus-reinforcer relations. *Journal of the Experimental Analysis of Behavior, 51,* 65–76.

Dymond, S., & Rehfeldt, R. A. (2000). Understanding complex behavior: The transformation of stimulus functions. *The Behavior Analyst, 23,* 239–254.

Edwards, C. A., Jagielo, J. A., Zentall, T. R., & Hogan, D. E. (1982). Acquired equivalence and distinctiveness in matching-to-sample by pigeons: Mediation by reinforcer-specific expectancies. *Journal of Experimental Psychology: Animal Behavior Processes, 8,* 244-259.

Galizio, M., Stewart, K. L., & Pilgrim, C. (2001). Clustering in artificial categories: An equivalence analysis. *Psychonomic Bulletin & Review, 8,* 609–614.

Gluck, M. A., & Bower, G. H. (1988). From conditioning to category learning: An adaptive network model. *Journal of Experimental Psychology: General, 8,* 37–50.

Hayes, S. C. (1989). Nonhumans have not yet shown stimulus equivalence. *Journal of the Experimental Analysis of Behavior, 51,* 385–392.

Hayes, S. C., Barnes-Holmes, D., & Roche, B. (2001). *Relational frame theory: A post-Skinnerian account of human language and cognition.* New York: Plenum.

Hayes, S. C., & Bisset, R. (1998). Derived stimulus relations produce mediated and episodic priming. *The Psychological Record, 48,* 617–630.

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

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

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

Honig, W. K., & Fetterman, J. G. (1992). *Cognitive aspects of stimulus control.* Hillsdale, NJ: Erlbaum.

Horne, P. J., & Lowe, C. F. (1996). On the origins of naming and other symbolic behavior. *Journal of the Experimental Analysis of Behavior, 65,* 185–241.

Kaiser, D. H., Sherburne, L. M., Steirn, J. N., & Zentall, T. R. (1997). Perceptual learning in pigeons: Decreased ability to discriminate samples mapped onto the same comparison in many-to-one matching. *Psychonomic Bulletin & Review, 4,* 378–381.

Kastak, C. R., Schusterman, R. J., & Kastak, D. (2001). Equivalence classification by California sea lions using class-specific reinforcers. *Journal of the Experimental Analysis of Behavior, 76,* 131–158.

Keller, F. S., & Schoenfeld, W. N. (1950.) *Principles of psychology.* New York: Appleton-Century-Crofts.

Lakoff, G. (1987). *Women, fire, and dangerous things: What categories reveal about the mind.* Chicago: University of Chicago Press.

Lane, S. D., Clow, J. K., Innis, A., & Critchfield, T. S. (1998). Generalization of crossmodal stimulus equivalence classes: Operant processes as components in human category formation. *Journal of the Experimental Analysis of Behavior, 70,* 267–279.

Lane, S. D., & Critchfield, T. S. (1998). Classification of vowels and consonants by individuals with moderate mental retardation: Development of arbitrary relations via match-to-sample training with compound stimuli. *Journal of Applied Behavior Analysis, 31,* 21–41.

Laurence, S., & Margolis, E. (1999). Concepts and cognitive science. In E. Margolis & S. Laurence (Eds.), *Concepts: Core readings* (pp. 3–81). Cambridge, MA: MIT Press.

Manabe, K., Kawashima, T., & Staddon, J. E. R. (1995). Differential vocalization in budgerigars: Towards an experimental analysis of naming. *Journal of the Experimental Analysis of Behavior, 63,* 111–126.

Margolis, E., & Laurence, S. (Eds.). (1999). *Concepts: Core readings.* Cambridge, MA: MIT Press.

Markman, E. M. (1991). The whole-object, taxonomic, and mutual exclusivity assumptions as initial constraints on word meanings. In S. A. Gelman & J. P.

Byrnes (Eds.), *Perspectives on language and thought* (pp. 72–106). New York: Cambridge University Press.

McDaniel, M. A., Nuefeld, K. H., & Damico-Nettleton, S. (2001). Many-to-one and one-to-many associative learning in a naturalistic task. *Journal of Experimental Psychology: Applied, 7(3),* 182–194.

Newman, F. L., & Baron, M. R. (1965). Stimulus generalization along the dimension of angularity. *Journal of Comparative and Physiological Psychology, 60,* 59–63.

Pilgrim, C., & Galizio, M. (1995). Reversal of baseline relations and stimulus equivalence: I. Adults. *Journal of the Experimental Analysis of Behavior, 63,* 225–238.

Rescorla, R. A., & Wagner, A. R. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. In A. H. Black & W. F. Prokasy (Eds.), *Classical conditioning: Vol. 2. Current research and theory* (pp. 64–99). New York: Appleton-Century-Crofts.

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

Rosch, E. (1978). Principles of categorization. In E. Rosch & B. B. Lloyd (Eds.), *Cognition and categorization* (pp. 27–48). Hillsdale, NJ: Erlbaum.

Saunders, R. R., & Green, G. (1999). A discrimination analysis of training-structure effects on stimulus equivalence outcomes. *Journal of the Experimental Analysis of Behavior, 72,* 117–137.

Schwartz, B., Wasserman, E. A., & Robbins, S. J. (2002). *Psychology of learning and behavior*. New York: Norton.

Sidman, M. (1971). Reading and auditory-visual equivalences. *Journal of Speech and Hearing Research, 14,* 5–13.

Sidman, M. (1994). *Equivalence relations and behavior: A research story.* Boston: Authors Cooperative.

Sidman, M. (2000). Equivalence relations and the reinforcement contingency. *Journal of the Experimental Analysis of Behavior, 74,* 127–146.

Sidman, M., Kirk, B., & Willson-Morris, M. (1985). Six-member stimulus classes generated by conditional-discrimination procedures. *Journal of the Experimental Analysis of Behavior, 43,* 21–42.

Sidman, M., Rauzin, R., Lazar, R., Cunningham, S., Tailby, W., & Carrigan, P. (1982). A search for symmetry in the conditional discrimination of rhesus monkeys, baboons, and children. *Journal of the Experimental Analysis of Behavior, 37,* 23–44.

Sidman, M., & Tailby, W. (1982). Conditional discrimination vs. matching to sample: An expansion of the testing paradigm. *Journal of the Experimental Analysis of Behavior, 37,* 5–22.

Skinner, B. F. (1945). The operational analysis of psychological terms. *Psychological Review, 52,* 270–277.

Skinner, B. F. (1977). Why I am not a cognitive psychologist. *Behaviorism, 5,* 1–10.

Smith, E. E., & Medin, D. L. (1981). *Categories and concepts.* Cambridge, MA: Harvard University Press.

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

Urcuioli, P. J., & Zentall, T. R. (1993). A test of comparison-stimulus substitutability following one-to-many matching by pigions. *The Psychological Record, 43,* 725–744.

Urcuioli, P. J., Zentall, T. R., Jackson-Smith, P., & Steirn, J. N. (1989). Evidence for common coding in many-to-one matching: Retention, intertrial interference, and transfer. *Journal of Experimental Psychology: Animal Behavior Processes, 15,* 264–273.

Vaughan, W., Jr. (1988). Formation of equivalence sets in pigeons. *Journal of Experimental Psychology: Animal Behavior Processes, 14,* 36–42.

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

Wilkinson, K. M., Dube, W. V., & McIlvane, W. J. (1996). A crossdisciplinary perspective on studies of rapid word mapping in psycholinguistics and behavior analysis. *Developmental Review, 16,* 125–148.

Wixted, J. T., & Gaitan, S. C. (in press). Cognitive theories of reinforcement history surrogates: The case of likelihood ratio models of human recognition memory. *Animal Learning & Behavior*.

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

Wynne, C. D. L. (1995). Reinforcement accounts for transitive inference performance. *Animal Learning & Behavior, 23,* 207–217.

Zentall, T. R., Clement, T. S., & Weaver, J. E. (in press). Symmetry training in pigeons can produce functional equivalences. *Psychonomic Bulletin & Review.*

Zentall, T. R., & Hogan, D. E. (1978). Same/different concept learning in the pigeon: The effect of negative instances and prior adaptation to the transfer stimuli. *Journal of the Experimental Analysis of Behavior, 30,* 177–186.

Zentall, T. R., Sherburne, L. M., & Urcuioli, P. J. (1993). Common coding in a many-to-one delayed matching task as evidenced by facilitation and interference effects. *Animal Learning & Behavior, 21,* 233–237.

Zentall, T. R., & Smeets, P. M. (Eds.). (1996). *Stimulus class formation in humans and animals.* Amsterdam: Elsevier.

Zentall, T. R., Steirn, J. N., Sherburne, L. M., & Urcuioli, P. J. (1991). Common coding in pigeons assessed through partial versus total reversals of many-to-one conditional and simple discriminations. *Journal of Experimental Psychology: Animal Behavior Processes, 17,* 194–201.

Received June 19, 2002
Final acceptance July 1, 2002