

Relational Frame Theory: A New Paradigm for the Analysis of Social Behavior

Bryan Roche, Yvonne Barnes-Holmes, Dermot Barnes-Holmes,
Ian Stewart, and Denis O’Hora
National University of Ireland, Maynooth

Recent developments in the analysis of derived relational responding, under the rubric of relational frame theory, have brought several complex language and cognitive phenomena within the empirical reach of the experimental analysis of behavior. The current paper provides an outline of relational frame theory as a new approach to the analysis of language, cognition, and complex behavior more generally. Relational frame theory, it is argued, also provides a suitable paradigm for the analysis of a wide variety of social behavior that is mediated by language. Recent empirical evidence and theoretical interpretations are provided in support of the relational frame approach to social behavior.

Key words: relational frame theory, relational frames, derived stimulus relations, stimulus equivalence, social behavior

Relational frame theory (RFT; Hayes, 1991; Hayes, Barnes-Holmes, & Roche, 2001) provides a naturalistic and functional-analytic account of language and cognitive phenomena. This theory provides a definition of language and cognitive events in terms of derived relational responding. Moreover, RFT attempts to specify the types of social interactions that lead to the emergence of language and cognitive skills in the first instance (see Barnes, 1994; Y. Barnes-Holmes, Barnes-Holmes, Roche, & Smeets, 2001; Hayes, Gifford, & Ruckstuhl, 1996). For the reader to appreciate the potential of RFT for the analysis of complex behavior in general, and social behavior in particular, we must first outline the basic features of this theory. Hav-

ing considered RFT as an approach to complex behavior, we will then demonstrate how RFT can be applied theoretically and empirically in the analysis of a variety of social behaviors.

RELATIONAL FRAME THEORY

Most published research on derived relational responding has focused on the familiar stimulus equivalence effect, which can be described essentially as follows. If a verbally able human subject is trained, in a matching-to-sample context, to match A to B and B to C, he or she will also likely match C to A and A to C without reinforcement (see Fields, Adams, Verhave, & Newman, 1990; Sidman, 1992). Relational frame theory extends the analysis of this effect by treating stimulus equivalence as just one instance of a variety of derived stimulus relations (Barnes, 1994; Hayes, 1991, 1994; Hayes & Hayes, 1989; Steele & Hayes, 1991). More specifically, several studies to date have provided empirical evidence that it is possible for human subjects to respond in accordance with relations other than equivalence, such as difference, opposition, more than, and less than (Dymond & Barnes, 1995, 1996; O’Hora, Roche, Barnes-Holmes, & Smeets, in press; Roche & Barnes, 1996, 1997; Roche, Barnes-Holmes,

An early version of this paper was presented at the Third International Congress on Behaviorism and the Behavioral Sciences, Seville, Spain, 1998. A later version was presented at the Association for Behavior Analysis annual convention, New Orleans, May 2001. The first author thanks Janet Dillon for her help and support and an anonymous reviewer for several helpful suggestions.

Requests for reprints may be sent to Bryan Roche, Department of Psychology, National University of Ireland, Maynooth, Maynooth, Co. Kildare, Ireland (e-mail: Bryan.T.Roche@may.ie).

Smeets, Barnes-Holmes, & McGeady, 2000; Steele & Hayes, 1991). Several further studies have attempted to identify relational histories necessary to establish contextual stimulus functions for derived relational responding in the first instance (e.g., Y. Barnes-Holmes, Barnes-Holmes, Roche, & Smeets, 2001; Lipkens, Hayes, & Hayes, 1993; see also Y. Barnes-Holmes, Hayes, Barnes-Holmes, & Roche, 2001).

Most RFT studies to date have attempted to demonstrate that derived relations can transform the functions of events. For instance, if a child learns at school that men and women are of the *opposite* sex (i.e., a relation of opposition between *men* and *women*) and further learns that men are strong (i.e., a relation of coordination between the terms *men* and *strong*), they may derive that women are weak. If the child then further learns that strength comes with age, they may derive that younger women are weaker than older women, and that younger women are even weaker still than older men. In effect, a relational network of trained and derived stimulus relations pertaining to gender has been established (see Roche & Barnes, 1997, for a relevant study).

Relational frame theory is a conservative account of derived relational activity. This theory adopts the basic position that derived relational responding is *generalized operant behavior*. From this perspective, the act of relating is a contextually controlled overarching response class that can occur with an infinite variety of stimulus topographies. Relational frame theory suggests that contextual control for relational responding is established for humans during early language interactions. For example, children are often presented with objects and asked to repeat their names. This can be described as see Object A, hear Name B, and say Name B. Children are also taught to identify objects when they hear the appropriate name. This can be described as hear Name B, then orient towards Object A. Initially, each object-word and word-object relation is explicitly

trained. However, when a child has been exposed to enough of this relational training, generalized relational responding may emerge. Suppose, for example, that a child with this history of naming is taught "This is your toy." Contextual cues (such as the word *is* and the context of the social interaction more generally) predict that if this object is a "toy" (Object A, Name B), a "toy" is this object (Name B, Object A). Consequently, the child may now identify the object when asked "Where is your toy?" in the absence of prior differential reinforcement for doing so. Thus, deriving relations is not genuinely novel but is a type of generalized operant behavior (see D. Barnes-Holmes & Barnes-Holmes, 2000, for an extended discussion).

Patterns of relational responding are brought under the control of contextual cues (e.g., the word *is*) through a simple process of differential reinforcement. That is, to begin, both directions of a relation are explicitly trained (e.g., "A is B" and "B is A" are both reinforced). It is only when a sufficient range of such exemplars has been trained that the relation can generalize to a novel set of stimuli. For instance, following exemplar training for bidirectional relations and training in the relation "X is Y," a child may derive "Y is X" without reinforcement. Other than the topography of the stimuli involved, there is nothing new in this latter derived performance that has not been reinforced during training. However, the trained response has generalized and is now applicable to any set of relata. When this occurs, the pattern of generalized derived responding is referred to as a *relational frame*. Thus, RFT simply suggests that the well-established concept of the operant can be extended (to relational responding) to explain one of the key generative features of human language.

Other types of stimulus relations, such as comparison, may also be explained in terms of generalized operant behavior. Consider, for example, a young child who is taught to respond

to a range of questions such as “Which cup has *more* milk?” or “Which box has *more* toys?” Given sufficient exposure to such questions and appropriate reinforcement for answering them correctly, the appropriate response to the cups may come under the control of cues in addition to the actual relative quantities of milk in the cups (e.g., the word *more*). For example, a dime is *worth more than* a nickel, even though a dime is actually smaller in size than a nickel. As responses are brought under the contextual control of the appropriate cues (e.g., the word *more*), they become arbitrarily applicable. In other words, the child may eventually answer questions regarding relative quantity correctly, even with novel varieties of stimuli and in contexts in which the appropriate response cannot be based on the formal properties of the stimuli involved (e.g., the relative value of coins). When this occurs, a relational frame of *comparison* has been established.

Relational frame theory can be used in the analysis of entire networks of relations. Relational networks help us to describe the patterns according to which functions transform in the presence of particular contextual cues. Imagine, for instance, that we train the following network of relations: Stimulus A *bigger than* Stimulus B, B *bigger than* C, C *same as* D (i.e., $A > B > C = D$). Now suppose that we establish an aversive stimulus function for D by pairing it with an electric shock. Based on the established relational network, we should expect the conditioned aversive stimulus function to emerge for C (i.e., because it is the same as D), but to be amplified for B (because it is bigger than C), and to be amplified further for A (because it is bigger than B and bigger than C). Relational frame theory refers to this effect as a *transformation of functions* to describe changes in the functions of the A and B stimuli in accordance with the relational network in which they participate. Indeed, several studies have demonstrated this type of complex derived transformation of

function effect with operant and respondent functions (see Dymond & Barnes, 1995, 1996; Roche & Barnes, 1997; Roche et al., 2000).

Before presenting our RFT analysis of social behavior, we will consider one final feature of RFT that is relevant in the current context. Specifically, RFT places a special emphasis on the role of verbal consistency in maintaining verbal behavior in the absence of explicit reinforcement. As we shall see in later sections, the reinforcing effects of verbal consistency play an important role in several social psychological phenomena. Interestingly, it was research on stimulus equivalence that first alerted researchers to the role of response consistency in maintaining derived relational responding (e.g., Bush, Sidman, & de Rose, 1989; Spradlin, Cotter, & Baxley, 1973). Sidman (1994) noted during this early research on stimulus equivalence that class-consistent test performances often develop gradually without any apparent change in contingencies. In attempting to explain this effect, Sidman pointed to the fact that every stimulus is a member of multiple classes, in addition to those designated by the experimenter. For instance, although Stimulus A1 may participate in a class with Stimulus B1 based on reinforced conditional discriminations, from the subject's perspective B1 may also participate in a class with A2 based on formal similarities (e.g., both B1 and A2 may contain four sharp edges). Thus, given B1 as a sample in a test for symmetry with A1 and A2 as comparisons, the subject may choose A1 on the basis of experimental contingencies or A2 on the basis of preexperimental contingencies. In effect, responding to A2 is “incorrect” from the point of view of the experimenter but is “correct” from the point of view of the subject. Given that equivalence tests are conducted in extinction, there are no differential consequences for one response or the other. However, because sample and comparison stimuli vary across trials, the only response

pattern consistently available across all trials is that established by the experimental contingencies (e.g., matching on the basis of four sharp edges is only rarely an option). According to Sidman, it is only when a sufficient number of trials have been presented that the one consistent basis for responding becomes evident and begins to control responses. Across these trials, inconsistent response patterns are washed out and the relation that is possible on every trial eventually becomes the consistent basis for choice (see Harrison & Green, 1990, for empirical evidence of this effect).

Relational frame theory goes one step further by suggesting that, given a protracted history of reinforced language interactions within the social community, a multitude of contexts should support ongoing derived relational responding in the absence of reinforcement. In essence, RFT predicts that correspondences in a relational network (i.e., verbal coherence) will function as a relatively powerful reinforcer for relational activity itself, and that this will be an important feature of the verbal behavior of most individuals. Indeed, several studies have shown that say-do correspondences can be reinforced (Baer, Detrich, & Weninger, 1988; Catania, Shimoff, & Matthews, 1987; de Freitas Ribeiro, 1989; Lloyd, 1980; Paniagua, 1985; Riegler & Baer, 1989). Moreover, children are taught to be consistent in what they say and do from the time they can control their social environments through speaking. There are even strong sanctions against adults for being inconsistent in their behavior (see Guerin, 1994, p. 159). Other researchers have examined the social processes involved in the maintenance of verbal consistency across verbal episodes (see Schauss, Chase, & Hawkins, 1997). It should not be surprising, therefore, if verbal consistency, and behavior-behavior consistency more generally, were to become a conditioned reinforcer for verbal behavior itself.

Although behavioral consistency may

be trained in even the simplest of organisms, a special variety of consistency is available to the verbal organism. More specifically, given a history of derived relational responding, a verbal organism may respond to consistencies across behavior and verbal episodes that are topographically dissimilar. For instance, a speaker may respond to two different statements made in different languages as having the same meaning. Furthermore, they may respond to the verbal coherence that is obtained across the base and target domains specified in analogies (e.g., "hand is to glove as foot is to shoe") and metaphors (e.g., "cats are dictators") in a way that would seem impossible for a nonverbal organism (see Stewart, Barnes-Holmes, Roche, & Smeets, 2001). From the RFT perspective, once an extensive network of relational frames is established and a history of reinforcement is provided for producing coherent relational networks (i.e., not contradicting oneself), coherence will serve as a continuously available reinforcer for derived relational responding. That is, language will become a self-sustaining process because the very product of relational responding (i.e., coherent relational networks) is itself a conditioned reinforcer for further relational activity. As we shall see subsequently, verbal coherence plays an important role in several aspects of social behavior.

Relational frame theory is relatively new to the field of behavior analysis. In a short period, however, RFT has contributed to the analysis of language, cognition, and complex behavior generally (see Hayes, Barnes-Holmes, & Roche, 2001). However, many research questions still remain regarding the precise nature of the verbal and nonverbal interactions that generate and maintain derived relational responding. For instance, researchers have attempted to explain why some derived relations can be altered more readily than others when baseline conditional discriminations are altered during equivalence training (i.e., symmetrical and

transitive relations; see Pilgrim & Galizio, 1995). Others have questioned the high failure rate in arbitrary matching-to-sample tasks in verbally able children (Pilgrim, Jackson, & Galizio, 2000) and the failure to establish equivalence classes with verbally sophisticated adult humans, particularly when large numbers of stimuli are involved (see Fields et al., 1990). Indeed the very stability of derived equivalence relations has been questioned (Pilgrim et al., 2000; Spradlin, Saunders, & Saunders, 1992). Despite these research questions, RFT has nevertheless brought researchers into empirical contact with a host of hitherto unexamined complex behavioral phenomena and generated a healthy body of research questions. Some studies, for instance, have used the RFT approach to study syntax and grammar as products of highly complex contextual control for relational responding (Cullinan, Barnes, Hampson, & Lyddy, 1994). Other studies have applied RFT in the experimental analysis of analogy, metaphor (Stewart, Barnes-Holmes, Hayes, & Lipkens, 2001), and rule following (Hayes & Hayes, 1989; O'Hora & Barnes-Holmes, 2001). Some further studies have applied the RFT analysis to behavior that extends beyond the ostensibly verbal, such as sexual categorization (Roche & Barnes, 1996), sexual arousal (Roche & Barnes, 1997, 1998; Roche et al., 2000), and attitude formation and change (Grey & Barnes, 1996; Roche, Barnes, & Smeets, 1997). The current paper is primarily concerned with this latter type of research.

In what follows, we will show how the relational frame approach to verbal events can provide a philosophically coherent and technologically advanced approach to discourse, social knowledge, and behavior that will move us towards greater prediction and influence in these domains.

AN RFT APPROACH TO SOCIAL BEHAVIOR

Language is essentially a social activity (Skinner, 1957). Only a social

community can provide the learning history necessary to use words appropriately and to produce derived relational responding. Insofar as RFT is an adequate approach to language development and maintenance, then, it must necessarily comment on the relation between language and social behavior. Of course, other behavioral approaches to social behavior are possible within the framework of Skinner's *Verbal Behavior* (1957; see Guerin, 1994) or the sizable body of research on rule-governed behavior (Kunkel, 1997; see also Schauss et al., 1997). There are several technical problems, however, with Skinner's approach to verbal behavior and rule governance (e.g., O'Hora & Barnes-Holmes, 2001). An adequate consideration of these problems would itself merit extensive discussion; one has recently been provided within the pages of this journal (see D. Barnes-Holmes, Barnes-Holmes, & Cullinan, 2000, for a full discussion of the relation between RFT and Skinner's account of verbal behavior; see also Hayes, Blackledge, & Barnes-Holmes, 2001; Hayes & Hayes, 1989). Suffice to say at this point that RFT does not reject the Skinnerian approach but rather builds upon it with multiple derived stimulus relations (see Chase & Danforth, 1991, and Guerin, 1994, for analyses of language that supplement Skinner's *Verbal Behavior* with the concept of stimulus equivalence). Thus, although the current RFT approach to social behavior is not explicitly Skinnerian, we remain mindful that Skinner's account of verbal behavior has provided the conceptual bedrock for RFT itself.

Analyses of social behavior have also been provided from cultural anthropological (e.g., Harris, 1977) and Kantorian perspectives (Kantor, 1982; see also Houmanfar, Hayes, & Fredericks, 2001). Such analyses, however, are presented in terms of broad social, moral, and survival contingencies and are not sufficiently detailed and psychological to promise the immediate prediction and influence of the social

behavior of individuals in a laboratory context or otherwise.

In what follows we will outline a series of social psychological phenomena that have become amenable to behavioral research in terms of relational frame theory. In several cases, we will provide empirical data to support the RFT approach to social behavior. Our aim here, however, is not to present an already completed analysis of social psychological phenomena, but to outline an ongoing agenda of research and to share with the reader our vision of a technologically sophisticated and conceptually coherent approach to social psychology.

Social Categorization and Group Cohesion

Social categorization is traditionally thought to underlie social identity (e.g., Tajfel & Turner, 1979). For instance, by categorizing oneself and others as in-group or out-group members, individuals create a social identity for themselves and may benefit from an enhanced boost in self-esteem (Tajfel, 1982). Group membership is beneficial for the individual because it prescribes modes of action and reduces ambiguity in social situations. An important question for the behavior analyst, however, is how precisely does group membership influence the verbal behavior of members in social settings? In answering this type of question, social psychologists have focused their attention on mediating cognitive processes (e.g., Hogg & Turner, 1987). From a behavioral perspective, however, what is required is direct influence over the categorization process itself. One group of behavioral researchers has attempted to do just this in the context of political struggles in Northern Ireland (Watt, Keenan, Barnes, & Cairns, 1991).

In the Watt et al. (1991) study, a simple stimulus-equivalence-type paradigm was used to take advantage of the fact that people in Northern Ireland often respond to each others' names as discriminative for their religious back-

grounds. During the study, Northern Irish and English subjects were trained to relate three Catholic names to three nonsense syllables, and subsequently to relate the three nonsense syllables to three traditionally Protestant symbols (i.e., the words *Union Jack*, *Orange Order*, and *Lambeg Drum*). During the test for derived relations, subjects were presented with Protestant symbols as samples; the comparisons were two of the Catholic names used during training and a novel Protestant name. All of the English subjects chose the Catholic name (related through equivalence to the Protestant symbols), but 12 of the 19 Northern Irish subjects chose the Protestant name in the presence of the Protestant symbols, thereby failing to form laboratory-induced equivalence relations.

It might be suggested that the social contingencies of reinforcement that operate in the Northern Ireland community were responsible for the nonequivalence responding of the 12 Northern Irish subjects. In other words, only the Northern Irish subjects had been socially trained to respond to Protestant symbols and Catholic names as belonging to socially exclusive categories. Watt et al. (1991) suggested that on the basis of prior social training, Northern Irish subjects who failed the equivalence test may have been responding in accordance with the relational frame of sameness (i.e., Protestant symbols with Protestant names), and those Northern Irish subjects who passed may have been responding in accordance with the relational frame of opposition (i.e., Protestant symbols with Catholic names).

Although RFT provides nomenclature for describing the relevant stimulus relations involved in social categorization (e.g., same, opposite), it remains an empirical matter to identify precisely the individual social learning histories required to produce the performances observed in the Watt et al. (1991) study. For instance, in a laboratory setting RFT can clearly predict and allow control of the emergence of

specific functions for given stimuli in different contexts with respect to different histories of relational activity (see Roche et al., 2000). However, it is another matter to speculate on probable preexperimental histories. The interpretation of the Watt et al. data in terms of relational frames is speculative, but it is nevertheless consistent with recent research that provides strong clues as to the broad *types* of natural language histories required for such derived relational responding to emerge (e.g., Y. Barnes-Holmes, Barnes-Holmes, Roche, & Smeets, 2001, in press; Lipkens et al., 1993). The Watt et al. research also seems to suggest that we may use this approach in "diagnostic" procedures to examine the types of verbal relations that operate in social discourse. Indeed, this was attempted in a further study (Roche & Barnes, 1996) with respect to the categorization of sexual terms.

In the Roche and Barnes (1996) study, male and female undergraduate students were first exposed to a relational pretraining procedure to establish directly the contextual control functions of same, opposite, and different with three abstract stimuli. A sexual categorization test was administered to see if subjects would spontaneously categorize the words *penis* and *vagina* with *dominate*, *submit*, and *forget* according to the relations specified by the three contextual cues. Results showed that when the same contextual cue was present, all subjects reliably matched *penis* with *dominate* and *vagina* with *submit*. However, in the presence of the opposite contextual cue, subjects consistently matched *penis* with *submit* and *vagina* with *dominate*. Finally, when presented with *penis* or *vagina* in the presence of the different contextual cue, subjects always chose *forget*. In effect, the relational test identified some rather persistent sexual categories in operation within the verbal community in which the subject socialized.

The findings of Roche and Barnes (1996) provide behavior-analytic sup-

port for the possibility that gender and gender relations are socially constructed through language (e.g., Gergen, 1988). Feminist critics and social constructionists alike have argued that the characteristics of masculinity and femininity (e.g., dominance and submission) are deeply embedded in social discourse concerning men and women (e.g., Bem, 1993). Others have argued that these linguistic constructions of gender both reflect and prescribe modes of social action (see Gergen, 1988). Roche and Barnes (1996) agree with these views and, moreover, developed an experimental methodology and an empirical analysis of gender in terms of RFT. In this way, RFT represents the beginnings of a technical behavioral analysis of the social construction of gender.

Group Cohesion

Social psychologists have identified several factors that appear to determine the cohesiveness of social groupings, such as attraction (see Hogg & Hains, 1996), similarity (Goethals & Darley, 1977), shared perceived threats (Lanzetta, 1955; Turner, Pratkanis, Probasco, & Leve, 1992), or shared values and norms (Cota, Evans, Dion, Kilik, & Longman, 1995; Zaccaro & McCoy, 1988). However, the processes that give rise to enhanced cohesion are not clearly identified. One group of behavioral researchers has conceptualized the cohesiveness of groups in terms of the reinforcing effects of membership and the punishing effects of group desertion (Cota et al., 1995). Reinforcement and punishment surely play an important role in group cohesion, but these ideas fall short of accommodating the clearly verbal nature of the reinforcing effects of group membership. Specifically, important reinforcing factors such as status, a sense of belonging, and self-esteem appear to be inherently verbal. Similarly, the punishing effects of abandoning a group are never contacted *directly* by those who remain within the group, and as such

must be responded to as verbally stated future consequences in order to function as punishers. Thus, it appears that both positive and negative consequences of group membership may be contacted indirectly through verbal processes (e.g., "I feel like I belong," "I would be lost without the group").

From an RFT perspective, group membership depends upon complex patterns of derived relational responding with respect to oneself and other group members. One important idea for RFT, therefore, is that individuals learn to respond verbally to their own behavior with evaluations, interpretations, and categorizations. As soon as we can interact verbally with ourselves (e.g., "I am a Catholic"), we have verbally constructed a "conceptualized self." Of course, a group can also be conceptualized in this manner. As soon as we respond verbally to ourselves in relation to those around us (e.g., "I am one of 600 million Catholics on the planet"), we have verbally constructed a "conceptualized group." Furthermore, because of the flexibility of derived relational responding, statements may be made about a group that bear little relevance to experienced facts about the group. For example, an individual may speak of his or her home football team as "the greatest in the world" despite their appalling track record on the field. In this case, the greatness of the group is one of its verbally constructed features.

The discrimination of shared features, such as values and beliefs, across group members may lead to a strengthening of the conceptualized group as a verbally constructed entity. In technical terms, such abstracted similarities can increase cohesion because they serve as contextual cues for frames of coordination or hierarchical class membership being applied to group members. For example, groups may adopt universal dress or conduct codes to increase the salience of shared characteristics as defining properties of group membership and to decrease the salience of the unique characteristics of

members (particularly those characteristics that might conflict with group membership).

One factor that is widely considered to have a powerful effect on group cohesiveness is the perception of a common threat to group members (see Turner et al., 1992). Specifically, research has shown that when groups are under a perceived threat, members tend to become more concerned for the welfare of the group as a whole. Relational frame theory uses the concept of the conceptualized group to explain this effect. Specifically, in a conceptualized group an individual shares features verbally with other members that are important defining features of that group, such as religion, ethnicity, and so on. Insofar as group membership is a relational matter, a threat to a group is also a threat to individual members, and vice versa. In technical terms, a threat to the conceptualized group may transfer through hierarchical relations (e.g., from leader to followers) or through relations of coordination (e.g., from one member to another) to each individual in that group. Given this derived threat to each individual, it is hardly surprising that group cohesion often increases.

One way in which groups maintain or enhance cohesion involves the alteration of reinforcing functions through the use of *formative* and *motivative augmentals*. Augmentals are statements that establish or alter reinforcing consequences but do not change the probability of reinforcement contingent on behavior. As such, they are the terms used in RFT for dealing with establishing stimuli (Michael, 1982) and function-altering stimuli (Schlinger & Blakely, 1987). The functions of consequences can be changed as soon as they participate in derived relations. Consider, for example, the *motivative augmental* statement "Wouldn't a hot juicy cheeseburger be delicious right now?" If this statement leads to burger eating, we cannot call it a discriminative stimulus: Eating burgers is as likely to be reinforced in the absence of

the statement as in its presence. Rather, the statement operates as a verbal establishing stimulus that alters the reinforcing functions of cheeseburgers by bringing the listener into direct contact with the perceptual functions of eating a cheeseburger (e.g., the listener may actually salivate). A formative augmental, on the other hand, establishes reinforcing or punishing consequences for a previously neutral event. An example might be "these slips are worth chances on money prizes." If the slips now function for the first time as a reinforcer, the statement was a formative augmental. The effect may then be expanded through a relational network. For example, when money is an existing reinforcer, the phrase "push the button to earn slips" may now become the functional equivalent of "push the button to earn money." The important point, then, is that formative augmentals can contribute to behavioral regulation even if the "new" consequences are never actually contacted by the listener.

Augmental control plays an important part in the cohesion of religious groups when powerful reinforcers available outside the group (e.g., sex) and powerful punishers delivered inside the group (e.g., self-flagellation, fasting) can weaken the loyalty of group members. More specifically, religious groups often speak of biological reinforcers such as sex, drugs, food, alcohol, and rest as undesirable. Such reinforcers can threaten group cohesion because they often control behavior more effectively than the verbally constructed distant consequences of religious rule following. To ameliorate the possibility of membership loss as a result of these, it is often thought necessary to establish negative functions for these activities so that they participate in frames of coordination with terms such as *bad*, *evil*, *morally weak*, *dangerous*, and the like. When the threat of desertion is great, group members will exert considerable effort in forbidding contact with outsiders and in verbally constructing relations

of opposition and difference between group members and their families or mainstream society in general (e.g., the Moonies, the Branch Davidians). For instance, nonmembers may be verbally coordinated with terms such as *the problem*, or placed in frames of opposition with terms such as *the chosen ones*. The establishment of such verbal relations may have the effect of rendering the functions of contact with family and other nonmembers as less reinforcing, and possibly even aversive, thereby protecting the group from dissolution.

Just as cohesive groups must sometimes arrange the transformation of reinforcing functions, it is sometimes necessary to transform the functions of some primary punishers such that they function as reinforcers. Violent gang initiation rites serve as an example. More specifically, gangs sometimes make full group membership conditional upon the completion of a physical endurance test, often involving physical beatings by other group members. The endurance of such rituals is relationally coordinated with such terms as *got what it takes* and *made of the right stuff*. Indeed, even following a traumatic initiation rite, further physical assaults endured by a member on the gang's behalf will often secure increased status for that individual. By confining the source of reinforcement to activity within the group and framing previous reinforcers as punishers and former punishers as reinforcers, a group can verbally construct a new picture of social reality in which good becomes bad and bad becomes good (see Wilder & Shapiro, 1991). In this way, it becomes more difficult for any member of the group to return to the mainstream culture.

Attitudes and Attitude Change

It is almost a cliché to suggest that the study of attitudes is a cornerstone of social psychology. Different researchers have defined attitudes in different ways, but most widely accepted

definitions make reference to an affective or evaluative component as well as to cognitive and behavioral components (see Fiske & Taylor, 1991, p. 463). One behavior-analytic study attempted to use a stimulus equivalence paradigm to analyze the evaluative component and the transfer of evaluative functions in accordance with derived relations. These researchers (Grey & Barnes, 1996) conceived an attitude as a network of derived and explicitly reinforced stimulus relations according to which the functions of events are transformed (e.g., a negative attitude towards condom use could be seen as responding in accordance with an equivalence relation between actual condoms and descriptive terms such as *unromantic* and *disgusting*; see also Moxon, Keenan, & Hine, 1993; Schauss et al., 1997; Watt et al., 1991).

Grey and Barnes (1996) used a matching-to-sample procedure to establish three three-member equivalence classes (A1, B1, C1; A2, B2, C2; A3, B3, C3) using nonsense syllables as stimuli. One member from two of these classes (A1 and A2) was then used to clearly label one of two videocassettes. The videos contained film clips of either a romantic or a religious nature. Having viewed the videos, subjects were required to categorize as *good* or *bad* four additional videocassettes labeled with one of each of the remaining nonsense syllables from the A1 and A2 classes (i.e., B1, C1, B2, C2). Results showed that these additional videocassettes were categorized according to the established equivalence classes, despite the fact that subjects had not been exposed to the contents of those tapes. In effect, the study demonstrated the transfer of an evaluative response from A1 and A2 to further stimuli in accordance with derived relations.

The Grey and Barnes (1996) study demonstrated that attitudes may be usefully conceived of in terms of derived stimulus relations and the transfer of functions. However, that study did not attempt to alter the relations or functions involved, thereby providing a

more complete experimental analogue of attitude formation and change. One further study (Roche et al., 1997) attempted to do just that.

In the first of four experiments, Roche et al. (1997) established sexual arousal functions for two nonsense syllables, which we will refer to here as A1 and C1, and nonsexual functions for two further nonsense syllables, which we will call A2 and C2. This was achieved by pairing presentations of A1 and C1 with sexual film clips and A2 and C2 with nonsexual film clips on a TV monitor. The functions of the stimuli were then assessed by the outcome of a conditional discrimination test presented to subjects on a computer. On this test, all 5 subjects matched stimuli on the basis of similarity in function (i.e., matched A1 with C1 and A2 with C2).

Subjects were then trained on the following conditional discriminations, A1-B1, A2-B2, B1-C2, B2-C1, after which the conditional discrimination test was again administered. Given this training, we would expect subjects to produce the equivalence relations A1-B1-C2 and A2-B2-C1. However, these equivalence relations would also contain members with incongruous functions (e.g., A1 and C2 are sexual and nonsexual, respectively). Results from Experiment 1 showed that all 5 subjects continued to match A1 with C1 and A2 with C2 in accordance with the original stimulus-pairing contingencies. Thus, it appeared that subjects who first matched stimuli on the basis of their functions following the pairing procedure were unlikely to produce the equivalence relations containing emotionally incongruous members (i.e., matching-to-sample training failed to override the sexual-nonsexual stimulus-pairing history).

In a further condition, subjects were first exposed to the conditional discrimination training. Having passed the equivalence tests, subjects were exposed to an incongruous stimulus-pairing procedure in which A1 and C2 were paired with sexual stimuli and A2

and C1 were paired with nonsexual stimuli. On a subsequent conditional discrimination test, all subjects continued to match stimuli based on derived equivalence relations (i.e., stimulus pairings failed to override the matching-to-sample training). In effect, Experiment 1 showed that once a response pattern was established using a stimulus pairing or conditional discrimination training procedure, this performance was persistent even after exposure to incongruous training phases that we might otherwise expect to give rise to alternative performances. Indeed, in a further experiment (Experiment 3), subjects were repeatedly exposed to these incongruous training phases without affecting their original conditional discrimination test performance.

Experiment 2 omitted the conditional discrimination test following the initial training phase (i.e., conditional discrimination or stimulus-pairing training). Results indicated that when the initial training was untested, subsequent conditional discrimination tests were controlled by the most recent training rather than by the initial training. In effect, persistent conditional discrimination test performances were not controlled by initial training conditions alone. Persistent performances arose only when subjects produced derived relations based on their initial training *before* being exposed to incongruous contingencies.

Experiment 4 tested the possibility that the incongruous contingencies failed to control behavior precisely because of their novel trial format (i.e., stimulus pairing vs. match to sample). In this experiment, subjects were first exposed to the types of contingencies that would later serve as incongruous. More specifically, some subjects were exposed to preliminary stimulus pairing and conditional discrimination testing with one set of stimuli before a second stimulus set was used to conduct conditional discrimination training and testing followed by incongruous stimulus pairing and testing. Other subjects

were exposed to preliminary conditional discrimination training and testing with one stimulus set before a second stimulus set was used to conduct stimulus pairing and testing followed by incongruous conditional discrimination training and testing. Results indicated that after a history of stimulus pairing and testing, equivalence-based test performances were often altered by the incongruous stimulus pairing and testing. Similarly, after a history of conditional discrimination training, stimulus-pairing-based test performance was often altered by incongruous conditional discrimination training and testing. Thus, a critical controlling variable for test performances appeared to have been the subjects' previous history of exposure to particular forms of stimulus control.

These analogue studies may represent the first steps towards an experimental analysis of attitude formation and change. Specifically, the findings suggest that once an attitude has been formed (e.g., an equivalence relation between *condoms* and *disgusting*), this relation may persist despite personal experiences that might be expected to undermine such attitudes (e.g., seeing a sexual health advertisement). Although this is an alarming finding, Roche et al. (1997) suggested a means by which we might disrupt such persistent attitudes.

Consider, once again, Experiment 4 in which some subjects were provided with a history of responding in accordance with conditional discriminations unrelated to those trained and tested subsequently. These subjects were then exposed to stimulus pairing and conditional discrimination testing. All subjects responded to the test by matching samples and comparisons on the basis of their common sexual and nonsexual functions. Finally, the subjects were exposed to incongruous conditional discrimination training, after which they *shifted* their stimulus-pairing-based performance to one based on derived equivalence relations. In effect, the pretraining appeared to sensitize

subjects' test performances to new learning experiences. Although this finding involved training manipulations made before the persistent test performances could emerge, the experiment suggests ways in which we might control already-established persistent attitudes.

As an example, consider a man who has heard and persistently believes that "real men don't use condoms." As the Roche et al. (1997) study suggests, the attempts of a community worker to undermine this attitude by repeatedly contradicting him, or by introducing him to individuals who have been exposed to the AIDS virus through unsafe sex, may do little to alter his attitude towards condom use. However, the results of Experiment 4 suggest that the negative attitude may persist because the individual has no history of reinforcement for responding in the context of the community worker. More specifically, subjects who were provided with the opportunity to produce or derive relations trained in a particular context (i.e., stimulus pairing or matching to sample) were more sensitive to deriving relations trained in that way in the future. This suggests two ways in which we might address a persistent attitude. First, we might expose this individual to a form of "novel pretraining," in which we familiarize him or her with a community worker as a general source of behavioral control. One way to gain behavioral control easily would be to establish novel verbal relations that are already consistent with verbal networks for the listener. For instance, if the individual in question is a sports enthusiast, a community worker might attempt to establish such relations as "regular exercise and a balanced diet prolong life." In effect, a variety of attitudes unrelated to the issue of condom use could be reinforced. Following a history of establishing and reinforcing verbal relations, the community worker might attempt to establish incongruous verbal relations such as "real men *do* wear condoms." In commonsense

terms, the community worker could first establish him or herself as a reliable source of information before attempting to shift persistent attitudes.

Second, the community worker might deliver a message through a person who has already demonstrated an influence over the individual's behavior. A respected and familiar person who has already demonstrated behavioral control should serve as an effective agent of attitude change when incongruous and novel verbal relations are presented. Indeed, this conclusion has been supported by intervention studies that have found that the safe-sex practices employed by eminent members of social groups were eventually emulated by members of those groups (Kelly et al., 1991, 1992; Wulfert & Biglan, 1994).

Of course, the RFT analysis of attitude formation and change is interpretive. More detailed empirical analyses may establish that the effectiveness of a familiar individual in delivering a novel message depends partly upon associative processes and a history of direct reinforcement for rule following. Nevertheless, it is almost inconceivable that attitude formation and change are entirely nonverbal processes that involve no derived relational activity. The analogue studies described here represent only the first step in the experimental analysis of those behavioral processes involved in the formation and maintenance of persistent attitudes. Moreover, we have only just begun to analyze methods that might be used for altering persistent attitudes when they are to the detriment of the individual or the wider culture. Nevertheless, the relatively consistent data reported by Roche et al. (1997) suggest that further basic research along the foregoing lines would be well worthwhile.

Persuasion and Rhetoric

The first widely held view of persuasion as a psychological process was known as the Yale model (Hovland, Lumsdaine, & Sheffield, 1949; Hov-

land & Mandell, 1952). According to this model, beliefs are swayed primarily by coherent arguments. Persuasion is most effective, therefore, when it appeals to logic. Hovland and his colleagues proposed that persuasion efforts must involve getting individuals to notice, understand, and remember an alternative viewpoint. They also argued that propaganda works by suppressing intellectual reasoning during the delivery of a novel message, for instance by bombarding an individual with information.

Recent RFT evidence shows, however, that once derived, verbal relations may resurge at a later time in an appropriate context without additional training designed to bring about such resurgence (Wilson & Hayes, 1996; see also Healy, Barnes-Holmes, & Smeets, 2000). From the RFT perspective, therefore, the propagandist might do well to alter the functions of already-established verbal relations through rhetoric rather than attempting to extinguish them (see also Roche et al., 1997).

Perhaps the simplest form of persuasion or rhetoric involves the alteration and weakening of psychological functions maintained by verbal relations for the listener. Weakening these functions is perhaps the most fundamental aspect of rhetoric because individuals are slow to respond positively to messages that compete with their beliefs and opinions (Hovland et al., 1949). As an example, consider a salesperson who must regularly contend with such problematic socially established verbal relations as "all salespeople are liars." Once a relational frame is established between such terms as *salesperson* and *liar*, it is difficult for the salesperson to say anything without having the functions of "liar" evoked for the listener. One rhetorical means by which the functions of the problematic relational network can be weakened, however, is to use terms and phrases that coordinate with trustworthy individuals and that participate in frames of opposition with dishonest salespersons. For ex-

ample, the salesperson might say, "you are under no obligation to buy," or "if you are not completely satisfied you can have your money back," or "take it now for free and if you like it you can send us the money next month." Thus, the pitch will participate in a relational network with other terms (e.g., *trustworthy*), stimuli (e.g., items purchased in a department store), and consequences (pleasure of using recently purchased items) that form part of any normal sales process. In effect, use of these and similar phrases may transform the functions of the sales pitch so that it does not evoke functions of dishonesty.

As suggested above, RFT maintains that a salesperson is unlikely to weaken problematic "dishonesty functions" by contradicting them directly. Specifically, direct reference to dishonesty itself is bound only to evoke the relevant functions, even when the term is placed in a frame of opposition with salespeople. In the same way, it is impossible to literally follow the rule "do not think of a pink elephant" because in order to do so, one must first respond covertly to the visual perceptual features of a pink elephant. Thus, an ineffective way for a salesperson to generate trust would be to use the phrase "trust me," because the phrase may actually serve to strengthen the problematic functions of dishonesty attached to salespeople.

Motivative augmentals may also be used to increase the value of verbally stated consequences. Health and life insurance companies, for example, rely heavily on this technique. Advertisements for such companies often aim to bring the listener into direct emotional contact with verbally constructed consequences of not purchasing insurance. Whereas these companies procure their profits on the basis that a greater number of people buy insurance than will need it, they nevertheless speak of these risks in such a way that they become psychologically proximate. This is achieved by framing aversive events in temporal relations. Temporal rela-

tions are based on the nonarbitrary experience of change, and like all relations can be arbitrarily applied to transform the functions of stated events (e.g., "old age is closer than you think"). For the same purpose, advertisers may also employ if-then relations (e.g., "if you injure your back you will not be able to support your family") and relations of coordination (e.g., "this could happen to you"). Such relations serve to transform the functions of the stated future consequences such that their reinforcing and punishing functions are enhanced. In each of these cases, of course, deictic relations are also involved. Deictic relations are those that specify a relation in terms of the perspective of the speaker, such as left-right, I-you, here-there, and now-then (see Barnes & Roche, 1997; Hayes, 1984). Thus, most cases of rhetoric require the listener to respond from a perspective different from that of the speaker. That is, a listener may respond to "this could happen to you" as "that could happen to me." (See Barnes & Roche, 1997; Y. Barnes-Holmes, Hayes, Barnes-Holmes, & Roche, 2001, for further information on the development of deictic relations.)

The foregoing RFT analysis is consistent with the rationale of the theory of reasoned action (Ajzen & Fishbein, 1980), which attempts to make sense of attitude and behavior change in terms of an individual's perception of susceptibility to aversive consequences (e.g., developing lung cancer because of smoking), their perceived competence in avoiding negative consequences (e.g., quitting smoking), and the value placed on the consequences of a behavior change (e.g., avoiding cancer). Thus, persuasive narratives that are emotionally provocative, informative, and directive should be effective in generating a change in attitude and an attendant change in behavior. From the RFT perspective, such narratives bring aversive verbally constructed consequences into psychological proximity. In addition, the provision of a directive

rule that evokes functions of escape from these aversive consequences through appropriate relational frames (e.g., "buy our insurance and sleep easy") may make rule following more likely.

One further technique commonly used by rhetoricians for the transformation of verbal functions involves the showcasing of a listener's verbal inconsistency. As outlined earlier, verbal consistency likely serves as a powerful reinforcer for verbal behavior itself. Moreover, verbal inconsistency is established as a powerful punisher for most members of the verbal community (see Guerin, 1994, p. 159; Schauss et al., 1997). Convincing speakers are skilled at using the listener's verbal inconsistency as a punisher. The effectiveness of the technique is based on the simple fact that nobody likes to feel foolish or confused. The rhetorician, therefore, will find ways to evoke the frustrating functions of two incongruent verbal relations that are produced by the listener. An insurance advertisement, for instance, may read, "People who care about their families are insured with Acme insurance." Here the listener must respond to a verbal relation (i.e., the slogan) that is not consistent with other relations in the listener's verbal repertoire (e.g., "I care about my family and I do not have Acme insurance"). If verbal consistency is to be maintained, then one of the two verbal relations must change. Interestingly, the problem faced by the listener in this case is similar to that presented to subjects in Festinger's (1957) cognitive dissonance studies. If effective, the rhetorical devices used by advertisers will be powerful enough that the verbal behavior of some individuals will be transformed such that consistency is achieved (i.e., the listener purchases Acme insurance). For other individuals, however, previously established verbal relations such as "advertisers will say anything to get you to buy their products" will reduce the impact of advertising slogans.

Conclusion

Relational frame theory represents a relatively recent advance within the evolution of behavior analysis. In a short time, RFT has shown itself to be a viable paradigm for the experimental investigation of complex human behavior (see Hayes, Barnes-Holmes, & Roche, 2001). In addition, the limited research reviewed here suggests that RFT may also provide a useful research route to many social psychological phenomena that hitherto lay beyond the empirical reach of behavior analysis. Of course, RFT was applied here somewhat speculatively to a range of social phenomena, and much work remains to be done in assessing its utility in this regard. Nevertheless, it has served us well in the current context by providing a vital new approach to a host of important psychological questions and generating a wealth of empirical research into a diverse range of psychological phenomena.

REFERENCES

- Ajzen, I., & Fishbein, M. (1980). *Understanding attitudes and predicting social behavior*. Englewood Cliffs, NJ: Prentice Hall.
- Baer, R. A., Detrich, R., & Weninger, J. M. (1988). On the functional role of the verbalization in correspondence training procedures. *Journal of Applied Behavior Analysis, 21*, 345–356.
- Barnes, D. (1994). Stimulus equivalence and relational frame theory. *The Psychological Record, 44*, 91–124.
- Barnes, D., & Roche, B. (1997). A behavior-analytic approach to behavioral reflexivity. *The Psychological Record, 47*, 543–572.
- Barnes-Holmes, D., & Barnes-Holmes, Y. (2000). Understanding complex human behavior: Two approaches to the concept of generalized operant classes. *The Psychological Record, 50*, 251–265.
- Barnes-Holmes, D., Barnes-Holmes, Y., & Cullinan, V. (2000). Relational frame theory and Skinner's *Verbal Behavior*: A possible synthesis. *The Behavior Analyst, 23*, 69–84.
- Barnes-Holmes, Y., Barnes-Holmes, D., Roche, B., & Smeets, P. M. (2001). Exemplar training and a derived transformation of function in accordance with symmetry. *The Psychological Record, 51*, 287–308.
- Barnes-Holmes, Y., Barnes-Holmes, D., Roche, B., & Smeets, P. M. (in press). Exemplar training and a derived transformation of function in accordance with symmetry: II. *The Psychological Record*.
- Barnes-Holmes, Y., Hayes, S. C., Barnes-Holmes, D., & Roche, B. (2001). Relational frame theory: A post-Skinnerian account of human language and cognition. In H. W. Reese & R. Kail (Eds.), *Advances in child development and behavior* (Vol. 28, pp. 101–138). San Diego, CA: Academic Press.
- Bem, S. L. (1993). *The lenses of gender*. New Haven, CT: Yale University Press.
- Bush, K. M., Sidman, M., & de Rose, T. (1989). Contextual control of emergent equivalence relations. *Journal of the Experimental Analysis of Behavior, 51*, 29–45.
- Catania, A. C., Shimoff, E., & Matthews, B. A. (1987). Correspondence between definitions and procedures: A reply to Stokes, Osnes, and Guevremont. *Journal of Applied Behavior Analysis, 20*, 401–404.
- Chase, P. N., & Danforth, J. S. (1991). The role of rules in concept learning. In L. J. Hayes & P. N. Chase (Eds.), *Dialogues on verbal behavior* (pp. 205–225). Reno, NV: Context Press.
- Cota, A. A., Evans, C. R., Dion, K. L., Kilik, L., & Longman, R. S. (1995). The structure of group cohesion. *Personality and Social Psychology Bulletin, 21*, 572–580.
- Cullinan, V., Barnes, D., Hampson, P. J., & Lyddy, F. (1994). The contribution of behaviour analysis to a connectionist model of basic syntax. In G. Orchard (Ed.), *Neural computing research and applications* (Part 2, pp. 33–45). Belfast, Northern Ireland: INNA.
- de Freitas Ribeiro, A. (1989). Correspondence in children's self-report: Tacting and manding aspects. *Journal of the Experimental Analysis of Behavior, 51*, 361–367.
- Dymond, S., & Barnes, D. (1995). A transformation of self-discrimination response functions through the arbitrarily applicable relations of sameness, more than, and less than. *Journal of the Experimental Analysis of Behavior, 64*, 163–184.
- Dymond, S., & Barnes, D. (1996). A transformation of self-discrimination response functions in accordance with the arbitrarily applicable relations of sameness and opposition. *The Psychological Record, 46*, 271–300.
- Festinger, L. (1957). *A theory of cognitive dissonance*. Stanford, CA: Stanford University Press.
- Fields, L., Adams, B. J., Verhave, T., & Newman, S. (1990). The effects of nodality on the formation of equivalence classes. *Journal of the Experimental Analysis of Behavior, 53*, 345–358.
- Fiske, S. T., & Taylor, S. E. (1991). *Social cognition* (2nd ed.). Singapore: McGraw-Hill.
- Gergen, M. M. (1988). Towards a feminist metatheory and methodology in the social sciences. In M. M. Gergen (Ed.), *Feminist thought and the structure of knowledge* (pp. 87–104). New York: New York University Press.

- Goethals, G. R., & Darley, J. (1977). Social comparison theory: An attributional approach. In J. M. Suls & R. L. Miller (Eds.), *Social comparison processes: Theoretical and empirical perspectives* (pp. 259–278). Washington, DC: Hemisphere.
- Grey, I., & Barnes, D. (1996). Stimulus equivalence and attitudes. *The Psychological Record*, 46, 243–270.
- Guerin, B. (1994). *Analyzing social behavior: Behavior analysis and the social sciences*. Reno, NV: Context Press.
- Harris, M. H. (1977). *Cannibals and kings*. New York: Vintage Books.
- Harrison, R. J., & Green, G. (1990). Development of conditional and equivalence relations without differential consequences. *Journal of the Experimental Analysis of Behavior*, 54, 225–237.
- Hayes, S. C. (1984). Making sense of spirituality. *Behaviorism*, 12, 99–110.
- Hayes, S. C. (1991). A relational control theory of stimulus equivalence. In L. J. Hayes & P. N. Chase (Eds.), *Dialogues on verbal behavior* (pp. 19–40). Reno, NV: Context Press.
- Hayes, S. C. (1994). Relational frame theory: A functional approach to verbal events. In S. C. Hayes, L. J. Hayes, O. Sato, & K. Ono (Eds.), *Behavior analysis of language and cognition* (pp. 9–30). Reno, NV: Context Press.
- Hayes, S. C., Barnes-Holmes, D., & Roche, B. (Eds.). (2001). *Relational frame theory: A post-Skinnerian account of human language and cognition*. New York: Plenum.
- Hayes, S. C., Blackledge, J. T., & Barnes-Holmes, D. (2001). Language and cognition: Constructing an alternative approach within the behavioral tradition. In S. C. Hayes, D. Barnes-Holmes, & B. Roche (Eds.), *Relational frame theory: A post-Skinnerian account of human language and cognition* (pp. 3–20). New York: Plenum.
- Hayes, S. C., Gifford, E. V., & Ruckstuhl, L. E. (1996). Relational frame theory and executive function: A behavioral approach. In G. R. Lyon & N. A. Krasnegor (Eds.), *Attention, memory, and executive function* (pp. 279–306). Baltimore: Paul H. Brookes.
- Hayes, S. C., & Hayes, L. J. (1989). The verbal action of the listener as the basis for rule-governance. In S. C. Hayes (Ed.), *Rule-governed behavior: Cognition, contingencies and instructional control* (pp. 153–190). New York: Plenum.
- Healy, O., Barnes-Holmes, D., & Smeets, P. M. (2000). Derived relational responding as generalized operant behavior. *Journal of the Experimental Analysis of Behavior*, 74, 207–227.
- Hogg, M. A., & Hains, S. C. (1996). Intergroup relations and group solidarity: Effects of group identification and social beliefs on depersonalized attraction. *Journal of Personality and Social Psychology*, 70, 295–309.
- Hogg, M. A., & Turner, J. C. (1987). Intergroup behaviour, self categorisation and the salience of social categories. *British Journal of Social Psychology*, 26, 325–340.
- Houmanfar, R., Hayes, L. J., & Fredericks, D. W. (2001). Religion and cultural survival. *The Psychological Record*, 51, 19–37.
- Hovland, C. I., Lumsdaine, A. A., & Sheffield, F. D. (1949). *Experiments on mass communication*. Princeton, NJ: Princeton University Press.
- Hovland, C. I., & Mandell, W. (1952). An experimental comparison of conclusion-drawing by the communicator and by the audience. *Journal of Abnormal and Social Psychology*, 47, 581–588.
- Kantor, J. R. (1982). *Cultural psychology*. Chicago: Principia Press.
- Kelly, J. A., St. Lawrence, J. S., Diaz, Y. E., Stevenson, L. Y., Hauth, A. C., Brasfield, T. L., et al. (1991). HIV risk behavior reduction following intervention with key opinion leaders of a population: An experimental community-level analysis. *American Journal of Public Health*, 81, 168–171.
- Kelly, J. A., St. Lawrence, J. S., Stevenson, L. Y., Hauth, A. C., Kalichman, S. C., Diaz, Y. E., et al. (1992). Community AIDS/HIV risk reduction: The effects of endorsements by popular people in three cities. *American Journal of Public Health*, 82, 1483–1489.
- Kunkel, J. H. (1997). The analysis of rule-governed behavior in social psychology. *The Psychological Record*, 47, 699–716.
- Lanzetta, J. T. (1955). Group behavior under stress. *Human Relations*, 8, 29–52.
- Lipkens, G., Hayes, S. C., & Hayes, L. J. (1993). Longitudinal study of derived stimulus relations in an infant. *Journal of Experimental Child Psychology*, 56, 201–239.
- Lloyd, K. E. (1980). Do as I say, not as I do. *New Zealand Psychologist*, 9, 1–8.
- Michael, J. (1982). Distinguishing between discriminative and motivational functions of stimuli. *Journal of the Experimental Analysis of Behavior*, 37, 149–155.
- Moxon, P. D., Keenan, M., & Hine, L. (1993). Gender role stereotyping and stimulus equivalence. *The Psychological Record*, 43, 381–394.
- O'Hara, D., & Barnes-Holmes, D. (2001). The referential nature of rules and instructions: A response to *Instructions, Rules, and Abstractions: A Misconstrued Relation* by Emilio Ribes-Iñesta. *Behavior and Philosophy*, 29, 21–25.
- O'Hara, D., Roche, B., Barnes-Holmes, D., & Smeets, P. M. (in press). Response latencies to multiple derived stimulus relations: Testing two predictions of relational frame theory. *The Psychological Record*.
- Paniagua, F. A. (1985). Development of self-care skills and helping behaviors in adolescents in a group home through correspondence training. *Journal of Behavior Therapy and Experimental Psychiatry*, 16, 237–244.
- Pilgrim, C., & Galizio, M. (1995). Reversal of baseline relations and stimulus equivalence I.

- Journal of the Experimental Analysis of Behavior*, 63, 225–238.
- Pilgrim, C., Jackson, J., & Galizio, M. (2000). Acquisition of arbitrary conditional discriminations by young normally developing children. *Journal of the Experimental Analysis of Behavior*, 73, 177–193.
- Riegler, H. C., & Baer, D. M. (1989). A developmental analysis of rule-following. In H. W. Reese (Ed.), *Advances in child development and behavior* (Vol. 21, pp. 191–219). New York: Academic Press.
- Roche, B., & Barnes, D. (1996). Arbitrarily applicable relational responding and sexual categorization: A critical test of the derived difference relation. *The Psychological Record*, 46, 451–475.
- Roche, B., & Barnes, D. (1997). A transformation of responsively conditioned stimulus functions in accordance with arbitrarily applicable relations. *Journal of the Experimental Analysis of Behavior*, 67, 275–301.
- Roche, B., & Barnes, D. (1998). The experimental analysis of human sexual arousal: Some recent developments. *The Behavior Analyst*, 21, 37–52.
- Roche, B., Barnes, D., & Smeets, P. M. (1997). Incongruous stimulus pairing and conditional discrimination training: Effects on relational responding. *Journal of the Experimental Analysis of Behavior*, 68, 143–160.
- Roche, B., Barnes-Holmes, D., Smeets, P. M., Barnes-Holmes, Y., & McGeady, S. (2000). Contextual control over the derived transformation of discriminative and sexual arousal functions. *The Psychological Record*, 50, 267–291.
- Schauss, S. L., Chase, P. N., & Hawkins, R. P. (1997). Environment-behavior relations, behavior therapy and the process of persuasion and attitude change. *Journal of Behavior Therapy and Experimental Psychiatry*, 28, 31–40.
- Schlinger, H., & Blakely, E. (1987). Function-altering effects of contingency-specifying stimuli. *The Behavior Analyst*, 10, 41–45.
- Sidman, M. (1992). Equivalence relations: Some basic considerations. In S. C. Hayes & L. J. Hayes (Eds.), *Understanding verbal relations* (pp. 15–27). Reno, NV: Context Press.
- Sidman, M. (1994). *Equivalence relations and behavior: A research story*. Boston: Authors Cooperative.
- Skinner, B. F. (1957). *Verbal behavior*. New York: Appleton-Century-Crofts.
- Spradlin, J. E., Cotter, V. W., & Baxley, N. (1973). Establishing a conditional discrimination without direct training: A study of transfer with retarded adolescents. *American Journal of Mental Deficiency*, 77, 556–566.
- Spradlin, J. E., Saunders, K. J., & Saunders, R. R. (1992). The stability of equivalence classes. In S. C. Hayes & L. J. Hayes (Eds.), *Understanding verbal relations* (pp. 29–43). Reno, NV: Context Press.
- Steele, D., & Hayes, S. C. (1991). Stimulus equivalence and arbitrarily applicable relational responding. *Journal of the Experimental Analysis of Behavior*, 56, 519–555.
- Stewart, I., Barnes-Holmes, D., Hayes, S. C., & Lipkens, R. (2001). Relations among relations: Analogy, metaphor, and allegory. In S. C. Hayes, D. Barnes-Holmes, & B. Roche (Eds.), *Relational frame theory: A post-Skinnerian account of human language and cognition* (pp. 73–86). New York: Plenum.
- Stewart, I., Barnes-Holmes, D., Roche, B., & Smeets, P. M. (2001). Generating derived relational networks via the abstraction of common physical properties: A possible model of analogical reasoning. *The Psychological Record*, 51, 381–408.
- Tajfel, H. (1982). *Social identity and intergroup relations*. Cambridge, MA: Cambridge University Press.
- Tajfel, H., & Turner, J. C. (1979). An integrative theory of intergroup conflict. In W. Austin & S. Worchel (Eds.), *The social psychology of intergroup relations* (pp. 33–47). Monterey, CA: Brooks/Cole.
- Turner, M. E., Pratkanis, A. R., Probasco, P., & Leve, C. (1992). Threat, cohesion, and group effectiveness: Testing a social identity maintenance perspective on groupthink. *Journal of Personality and Social Psychology*, 63, 781–796.
- Watt, A., Keenan, M., Barnes, D., & Cairns, E. (1991). Social categorization and stimulus equivalence. *The Psychological Record*, 41, 33–50.
- Wilder, D. A., & Shapiro, P. (1991). Facilitation of outgroup stereotypes by enhanced ingroup identity. *Journal of Experimental Social Psychology*, 27, 431–452.
- Wilson, K. G., & Hayes, S. C. (1996). Resurgence of derived stimulus relations. *Journal of the Experimental Analysis of Behavior*, 66, 267–281.
- Wulfert, E., & Biglan, A. (1994). A contextual approach to research on AIDS prevention. *The Behavior Analyst*, 17, 353–363.
- Zacarro, S. J., & McCoy, M. C. (1988). The effects of task and interpersonal cohesiveness on performance of a disjunctive group task. *Journal of Applied Social Psychology*, 75, 268–273.