Hughes, S., & Barnes-Holmes, D. (2016). Relational Frame Theory: The Basic Account. In S. Hayes, D. Barnes-Holmes...

Chapter · April 2016

Sean Hughes
Ghent University
42 PUBLICATIONS **395** CITATIONS

Hughes, S., & Barnes-Holmes, D. (in press). Relational Frame Theory: The Basic Account. In S. Hayes, D. Barnes-Holmes, R. Zettle, and T. Biglan (Eds.), Handbook of Contextual Behavioral Science. New York: Wiley-Blackwell.

Relational Frame Theory: The Basic Account

Sean Hughes[12] and Dermot Barnes-Holmes[1]

National University of Ireland, Maynooth[1]
Ghent University[2]

Introduction

Imagine that you are a relative newcomer to the world of psychological science and that you have managed to find an Archimedean point from which to survey the contemporary landscape of the discipline. Taking a look around, you will observe a rich, vibrant and active country fragmented into a series of “intellectual communities” or sub-disciplines such as health, social and cognitive psychology, not to mention clinical, personality, and neuropsychology. Although you may notice that individuals are increasingly forging new connections with their counterparts from other areas, more often than not, these communities are interested in their own sets of questions and independently engaged in the development of their own methods and theories. One of the consequences of this fragmented approach is a massive proliferation of competing theories and models about highly specific phenomena that often appeal to radically different concepts or “analytic units.” At the same time, any attempt or interest in the development of “overarching” theories that cannot only account for highly specific events, but also connect entire sub-disciplines, has reduced to a trickle (although see Anderson, 2013; Garcia-Marques & Ferreira, 2011; Posner & Rothbart, 2007 for recent attempts). Thus, unlike the biological sciences, we have no widely-accepted theory like natural selection that applies to, and binds, seemingly unrelated areas (e.g., language, cognition, and emotion) in a relatively coherent or parsimonious way. Nor do we have overarching accounts such as Newtonian or quantum mechanics, that when combined, enable us to predict a wide range of outcomes, from the actions of a single individual to the behavior of entire groups or societies. No periodic table has emerged that specifies the basic psychological “units” of analysis, how these units relate to one another, or accommodate the movement from simple to increasingly complex behaviors. In short, psychology appears to be more a collection of “loosely related study areas than a coherent, unified and evolving science" (Yanchar & Slife, 1997, p.235).

Interestingly, and parallel to these developments, citizens of another intellectual country known as contextual behavioral science (CBS) have also sought to understand human language and cognition. Drawing on nearly a half century's worth of empirical findings, they have identified what they believe to be the core functional "unit" from which the rich diversity of human psychological life springs forth. Even more surprisingly, a rising tide of scientific studies indicate that this basic unit (termed arbitrarily applicable relational responding or AARR) allows for a whole host of complex behaviors to be predicted and influenced with precision, scope, and depth. These findings have led to the development and subsequent refinement of a functional-contextual account of human language and cognition known as relational frame theory (RFT; Hayes, Barnes, & Roche, 2001). Unlike many other theoretical enterprises in modern psychology (which tend to focus on specific features or aspects of a relevant domain), RFT operates with a relatively ambitious and extremely broad goal in mind: to develop an inductive, monistic, and functionally rooted account of language and cognition that can speak to topics as diverse as the origins of language and the emergence of self, to factors responsible for human suffering, intelligence, reasoning, and evaluation. One need only thumb through the pages of this book to see how RFT has brought new insights, and also new controversies, wherever it has led.

Over the next two chapters we hope to provide an accessible introduction to the foundations, nature, and implications of this new theory. We will illustrate how a deceptively simple idea (that AARR is a learned operant acquired early on in our development) has transformed our ability to predict and influence many complex human behaviors. Indeed, within the space of two decades, RFT researchers have linked relational responding to the development of language, reasoning, and inference, as well as self and perspective taking, implicit cognition, developmental disorders, psychopathology, intelligence, and organizational behavior. These basic findings have been used to inform progress in applied areas such as health psychology, clinical psychology, social psychology, consumer psychology, and neuropsychology. They have also led to the development of programs for teaching and remediating linguistic/cognitive deficits, directly informed the treatment of psychopathology, influenced how we approach the behavior of organizations, and stimulated new connections with cognitive and evolutionary science.

Given the sheer scope of what RFT sets out to achieve and the explosion of research and theorizing that has taken place over the last two decades, we have had to divide our story into two halves. In this chapter our main aim is to introduce the reader to the origins of, as well as arguments and evidence for, RFT. In Part I we trace the study of AARR to its historical roots and explain why this phenomenon has occupied the attention of behavioral scientists for over 40 years now. This section will also provide a general introduction to RFT and highlight how the ability to frame events relationally unlocks an incredible degree of flexibility when adapting to the physical and social world around us. In Part II we take a closer look at some of the core assumptions that underpin this account. We explain precisely what RFT researchers mean when they claim that AARR is a learned operant behavior and chart how this ability is acquired early in infancy and rapidly scales in complexity. Part III examines the main "families" or types of relational responses that have been empirically examined to date and discusses their respective similarities and differences, while Part IV demonstrates how the ability to AARR develops over time, is amenable to change, and falls under different forms of stimulus control. In the final section (Part V) we highlight a number of features of relational responding that will become important when linking RFT to language and cognition later on. Although RFT remains the subject of continued debate, we believe that it provides an important theoretical and empirical advance for behavior analysis in particular and scientific psychology more generally.

This basic treatment of the theory will provide the necessary foundation for much of what is discussed in Chapter Y. In the second half of our story we take the reader on a journey through the RFT literature, stopping to consider some of the key empirical and conceptual developments that have shaped our understanding of human language and cognition. Doing so will reveal how RFT has stimulated a rich, vibrant, and progressive program of research, generated a host of new procedures, and raised novel questions in the process. Readers who are primarily interested in how RFT has been interfaced with specific aspects of psychological science might benefit from proceeding directly to the next chapter. However, those looking for a more technical understanding of this account and to better appreciate its origins, assumptions, and aims should begin their journey here.

Part I: Background to the Development of Relational Frame Theory

Throughout much of the past century, the question of what makes humans unique has occupied considerable attention within the behavioral sciences. In the behavior analytic tradition, for example, it was assumed that those learning principles identified in nonhumans could stretch to, and account for, much of complex human behavior (see Dymond, Roche, & Barnes-Holmes, 2003; Hayes, 1987; Hayes, Barnes-Holmes, & Roche, 2001). This “continuity assumption” served as an “intellectual rudder” and guided much work in the field, with researchers focusing on nonhumans in order to identify general learning principles that could predict-and-influence the behavior of our own species. In many respects, this bottom-up strategy was a successful one, yielding concepts that apply equally to humans and nonhumans alike (e.g., reinforcement, punishment, generalization, discrimination, extinction, recovery, and habituation). However, when researchers turned their attention to complex human behavior, a number of important findings emerged, findings that hinted at learning processes or principles that may be unique to, or largely elaborated in, some species relative to others. Indeed, early evidence from three different research domains highlighted that humans consistently respond in ways that are not readily observed elsewhere in the animal kingdom.

Language

Surprisingly, and unlike much of nonhuman behavior, language refused to submit to an analysis in direct contingency terms, and attempts to do so faced numerous conceptual and empirical problems. For instance, Skinner (1957) devised a direct contingency account that defined verbal behavior as that which is "*reinforced through the mediation of other persons*" (p. 2), and "*where the 'listener' must be responding in ways that have been conditioned precisely in order to reinforce the behavior of the speaker*" (p. 225). However, this definition and the interpretive analysis that it occasioned were criticized on several grounds (see Gross & Fox, 2009; Hayes, Barnes-Holmes, & Roche, 2001; Leigland, 1997). Referring to verbal behavior as that which is reinforced via social mediation turned out to be too general a statement and one that made it difficult to distinguish verbal from any other social behavior. For example, according to Stewart and Roche (2013);

> In an operant experiment, for example, the behavior of the organism under investigation is reinforced by an experimenter who has been explicitly trained to do so. Thus, by Skinner's definition, the behavior of nonhumans in these experiments qualifies as verbal. (p.52)

In other words, given that organisms were already engaging in verbal behavior in the laboratory, it proved difficult to isolate the latter in order to study it. The above definition was also argued to be nonfunctional because it was based on the history of a second organism (the listener) rather than the organism of interest (the speaker). This introduced a scenario whereby the history of listener needed to be studied in order to understand the behavior of the speaker, and in no other area of behavioral thinking were functional response classes defined in this way. Rather behavior was (and still is) defined as a function of an organism's learning history and current contextual factors. The above definition also led to a paradoxical situation in which the behavior of the speaker was considered to be verbal, while the behavior of the listener was not, downplaying the importance of verbal comprehension relative to production. Finally, a number of authors pointed out that children learn thousands of words as well as a variety of "linguistic rules" that are often combined in the absence of direct instruction or experience (Chomsky, 1959; for a discussion see D. Barnes-Holmes & Murphy, 2007). Thus it seemed that a direct contingency approach failed to account adequately for two of language's core properties (generativity and flexibility).

While Skinner's analysis stimulated a number of empirical and practical applications in the domain of developmental disabilities, its volume and scope over the past 50 years has been limited (e.g., Dymond & Alonso-Álvarez, 2010; although see Greer, 2008; Sautter & LeBlanc, 2006; Schlinger, 2008, 2010). This interpretive analysis failed to equip researchers with a means to predict and influence a comprehensive range of verbal behaviors and did not translate into a rising cycle of research and analysis capable of stimulating new and important empirical questions about language itself.

Rule-Following

Lines of fracture between human and nonhuman behavior started to emerge elsewhere as well. A growing body of work on rule-following revealed that humans and nonhumans adapt to the same set of environmental regularities in dramatically different ways. Much of this research employed intermittent schedules of reinforcement wherein an organism was exposed to a learning task that sometimes reinforced high and (at other times) low rates of responding (Baron & Galizio, 1983; Lowe, 1979; Shimoff, Catania, & Matthews, 1981). Although nonhumans successfully completed such tasks, they often adjusted to contingencies in ways that differed to their human counterparts. Evidence suggested that these interspecies differences were due to the deployment of (covert) self-generated rules on the part of humans and that the effects of these rules could be (a) augmented when they were explicitly stated prior to the task, or (b) eliminated when steps were taken to minimize their impact on subsequent performance (see Hayes, Brownstein, Zettle, Rosenfarb, & Korn, 1986). Research on human operant behavior also revealed that people become insensitive to subsequent changes in the environment once their behavior comes under instructional control (Hayes, Brownstein, Haas, & Greenway, 1986), while developmental studies showed that preverbal infants respond in strikingly similar ways to nonhumans (unlike more verbally sophisticated children who respond in ways that mirror adults; Bentall, Lowe, & Beasty, 1985; Vaughan, 1985).

Overall, this work led to two important conclusions. The first was that humans frequently formulate verbal rules about contingencies in the wider world and that these rules are deployed in order to regulate how they respond to those contingencies (i.e., humans were interacting with the world through a “verbal lens”; Hayes, 1989b). The second was that - like language - rule-governed behavior stubbornly refused to be analyzed in direct contingency terms (see D. Barnes-Holmes, O’Hora, et al., 2001; O’Hora & Barnes-Holmes, 2004). For instance, Skinner (1969) suggested that rules or instructions could be defined as discriminative stimuli that “specify” a contingency between antecedents, responses, and consequences. Thus when a father gives his son the following instruction, “*Clean your room now and I will give you some pocket money later*,” the son is likely to tidy his room because a contingency has been specified between cleaning and receiving a reward. The problem for Skinner was this definition failed to clarify how rules or instructions come to function as “contingency specifying stimuli,” especially when the individual has never received direct training for following such instructions in the past. Nor did it explain why rules or instructions come to act “as though” they were discriminative stimuli in the absence of an appropriate history of differential reinforcement (Schlinger & Blakely, 1987).

To illustrate, imagine that a friend hears that you intend to vacation in Europe next summer and remarks, “*When you visit Paris make sure to climb the Eiffel Tower.*” Schlinger (1993) correctly pointed out that it is (a) the act of visiting Paris, and not this statement that evokes or sets the occasion for climbing the Eiffel Tower, and that (b) this climbing response was not established in the same way as other discriminative stimuli. Although Schlinger offered a more accurate description of instructions as “function-altering stimuli,” he did not outline the history of reinforcement that is necessary for stimuli to alter the operant and/or respondent properties of other stimuli. In short, the rule-governed behavior literature seemed to suggest that a key feature was missing from a sophisticated functional analysis of such behavior. It would only be with the discovery of derived stimulus relating (O’Hora, Barnes-Holmes, Roche, & Smeets, 2004) that researchers would be able to articulate how stimuli acquire the “specifying” or “function-altering” properties of instructions. It is to this topic that we now turn.

Stimulus Equivalence

The discovery of a phenomenon known as stimulus equivalence further cemented the idea that humans were able to respond in ways that could not be explained in direct contingency terms. Although the concept of equivalence had long attracted the attention of philosophers (e.g., Aristotle in De Memoria et Reminiscentia, trans. 1941) and behavioral researchers (e.g., Hull, 1934; Jenkins & Palermo, 1964), it was only after a pioneering set of studies by Murray Sidman in the 1970s that this “symbolic” type of behavior was subject to careful and systematic scrutiny (Sidman, 1971, 2000, 2009). Interest in this phenomenon stemmed from a rather puzzling finding: When participants are exposed to a series of conditional discriminations, the stimuli involved in those discriminations are spontaneously related in ways that were never directly trained or instructed. Consider an early study designed to teach a group of institutionalized teenage boys with severe developmental disabilities how to read (Sidman, 1971). The author found that when the boys were taught to select pictures (B) in the presence of certain spoken words (A) and to select written words (C) in the presence of those same spoken words (A) they did something entirely unexpected. Although they had never been trained to relate written words (C) and pictures (B), they could now do so. That is, they not only showed evidence of having learned the directly trained relations between A and B as well as A and C, but also responded in a number of untrained ways that traditional learning theories could not readily explain (e.g., they related B to C). Sidman proposed that this type of training caused the written words, spoken words, and pictures to become interchangeable or equivalent to one another, and as a result, he labeled this effect “stimulus equivalence.”

Researchers quickly realized that stimulus equivalence represented “something new” – a type of behavior in which humans could respond to stimuli and events as if they were related in the absence of any direct reinforcement or instruction. Even more interesting was the fact that these “*derived*” or “*emergent*” relations between stimuli were entirely unexpected and should not have occurred according to direct contingency accounts. For instance, an explanation of equivalence in terms of stimulus generalization seemed problematic given that the stimuli involved in those relations - printed words, spoken words, and pictures - bore no physical resemblance to one another. Nor could such outcomes be explained away as respondents or operants because they emerged in the absence of such a history of learning. Rather, it seemed as if stimuli had become “symbols” that were mutually substitutable, even though (in many cases) they had never been paired with, or directly related to, each other in the past. These early assumptions proved to be accurate and over the ensuing four decades an extensive literature emerged suggesting that stimulus equivalence could be obtained with a wide variety of stimuli, populations and procedures (for a review see Sidman, 1994, 2000, 2009). This work also revealed that when pigeons, rats, chimpanzees, and baboons were exposed to training procedures like that outlined above they consistently failed to produce outcomes similar to those observed in their human counterparts (Dugdale & Lowe, 2000; Brino, Campos, Galvão, & McIlvane, in press; Lionello-DeNolf, 2009; although see Hughes & Barnes-Holmes, 2014; Zentall, Wasserman, & Urcuioli, 2014). Thus stimulus equivalence seemed to represent a type of behavior that was highly elaborated in, or unique to, some species and absent in others.

Summary

The difficulty in accounting for language in direct contingency terms, combined with the striking difference between human and nonhuman operant behavior and the discovery of stimulus equivalence, led many researchers to the same conclusion: While most organisms adapt to the environment via direct contact with contingencies, humans appear to respond in ways that are inherently symbolic, flexible, and generative. Yet this finding introduced an additional set of questions: What type of learning history do people need in order to act as if stimuli are related in the absence of reinforcement or instruction? How and when do these relational abilities emerge and what role do they play in basic human language, rule-following, and stimulus equivalence? Could they also play a role in other psychological phenomena such as perspective-taking and self, analogical reasoning, as well as “fast” and “slow” cognition? In order to provide an answer to these questions, we first need to delve a little deeper into research on stimulus equivalence. As we shall see, work in this area set the stage for experimental methodologies and conceptual insights that would lead to a better understanding of derived stimulus relating - and by implication - human language and cognition.

Stimulus Equivalence: An Overview

According to Sidman (2000), in order for a behavior to qualify as an instance of stimulus equivalence, it must first demonstrate three core properties, which he termed (a) reflexivity, (b) symmetry, and (c) transitivity. The most basic of these is *reflexivity* which refers to the fact that - within an equivalence relation - each stimulus must be conditionally related to itself. For example, people should select spoken words (A) in the presence of spoken words (A), written words (B) in the presence of written words (B), and pictures (C) in the presence of pictures (C). *Symmetry* requires that the relationship between stimuli also be reversible, so that when a person is taught to select the written word D-O-G (B) in the presence of a spoken word “*DOG*” (A) he or she will also select the spoken word (A) in the presence of the written word (B) (i.e., responding as if “A is the same as B” leads to the derived relation “B is the same as A”). *Transitivity* refers to the fact that when two or more relations are trained, a novel set of derived relations also tend to emerge. Thus if people are taught to choose the written word D-O-G (B) in the presence of a spoken word “*DOG”* (A) and a picture of a dog (C) in the presence of the spoken word (A), a novel relation between the written word (B) and picture (C) will subsequently emerge. It is worth noting that this example actually involves combined symmetry and transitivity, which has been interpreted as providing a simple or abbreviated test for an equivalence relation (see Figure 1). A final feature of stimulus equivalence is the transfer of stimulus functions. Many researchers have found that when a function is explicitly trained to one member of an equivalence class, that same function may then transfer to the other members of the equivalence class without further training. For instance, if a fear-eliciting function is established for an actual dog (through the receipt of a bite) a child may come to respond with fear whenever they are presented with the written word D-O-G or the spoken word *"DOG"* (for related findings see Dougher, Perkins, Greenway, Koons, & Chiasson, 2002; Rodríguez-Valverde, Luciano, & Barnes-Holmes, 2009; Smyth, Barnes-Holmes, & Forsyth, 2006).

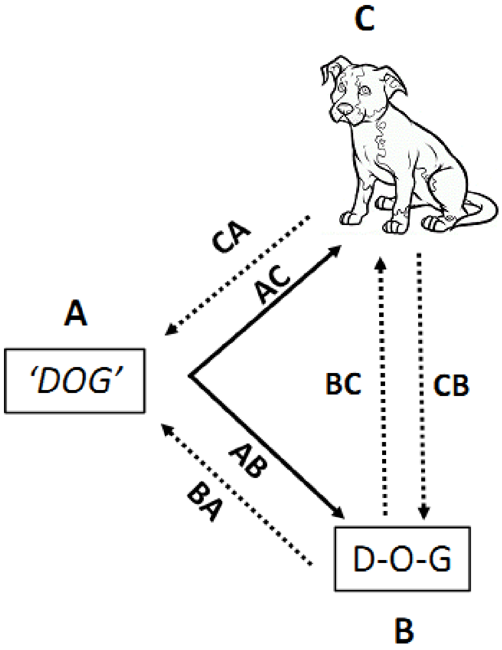

Figure 1. A visual illustration of an equivalence relation between the spoken word "*DOG"* (A), the written word D-O-G (B) and a picture of a dog (C). The solid arrows (AB and AC) designate relations between stimuli that are explicitly taught while the dashed arrows (BC and CB) indicate derived relations that emerge without any training or instruction. Note that testing only the B-C and C-B relations has sometimes been used as an abbreviated method for assessing equivalence responding (e.g. Devany, Hayes, & Nelson, 1986).

While stimulus equivalence generated a considerable amount of empirical and theoretical interest in its own right, it also set the stage for an entirely new possibility: If humans can derive that arbitrary stimuli such as spoken words, written words, and pictures are the same, then can they also derive other types of relations as well? In other words, are humans capable of responding in even more complex ways that extend above and beyond equivalence? If so, would these other types of *derived stimulus relations* also be reflexive, symmetrical, and transitive, and lead to transfers of function? After nearly three decades of work we now know that most humans are capable of deriving arbitrary relations among stimulus events without direct training or instruction to do so. We also know that equivalence is just the tip of the iceberg and that people can derive relations between stimuli in a near infinite number of ways. For instance, stimuli can be related as the same (e.g., "*Hond is the same as dog;*" Cahill et al., 2007) or opposite to one another (e.g., "*Good is the opposite of Evil;*"; Dymond, Roche, Forsyth, Whelan, & Rhoden, 2008), as well as hierarchically ("*Cat is a type of mamma;*" Gil, Luciano, Ruiz, & Valdivia-Salas, 2012), comparatively ("*Fruit is better than candy*;" Vitale, Barnes-Holmes, Barnes-Holmes, & Campbell, 2008), deictically ("*I am not you*;" McHugh & Stewart, 2012), and temporally ("*March comes before May*;" O'Hora et al., 2008) related. While these findings propelled our understanding of human learning forward, they also introduced an unavoidable conceptual problem: Many relational responses are difficult, if not impossible, to describe using the terms that Sidman originally devised in the context of stimulus equivalence.

Comparative relations, for example, are not symmetrical. If an elephant is bigger than a mouse, it does not follow that a mouse is bigger than an elephant (in fact, a mouse is smaller than an elephant). The same goes for causal relations. If "*Smoking causes cancer*" and "*Cancer causes death*," only some, but not all, of the properties of equivalence apply. Transitivity applies in that smoking causes death, but symmetry does not (given that cancer does not cause smoking).

Thus causal (and many other types of derived) relations are nonreflexive, asymmetrical, transitive, and connected. In other words, it quickly became apparent that humans were capable of relating stimuli in a vast number of different ways and that a new set of terms was needed that could adequately account for all possible derived relations that might emerge between and among stimuli. These technical terms would need to be broad enough so that they could not only describe the effects observed within the equivalence literature, but also encompass the properties of any other type of relation as well. Towards this end, researchers from a theoretical background known as relational frame theory have identified what they believe to be a small, but powerful set of terms that meet these various requirements. In the following section, we provide a general introduction to this account and consider these technical terms in greater detail.

Relational Frame Theory: An Overview

At its core, RFT argues that language, rule-following, and stimulus equivalence are all instances of a type of operant behavior known as arbitrarily applicable relational responding (AARR; D. Barnes-Holmes, Luciano, & Barnes-Holmes, 2004a, b; Dymond & Roche, 2013; Hayes, Barnes-Holmes, & Roche, 2001; Rehfeldt & Barnes-Holmes, 2009). According to this perspective, ‘relating” is a type of behavior and involves *responding to one event in terms of another*. While nonhumans and humans can both respond relationally to stimuli and events, the latter rapidly develop a more complex type of behavior (AARR) that fundamentally alters how they interact with the world around them. In what follows, we examine how RFT carves this type of operant behavior into two different varieties (nonarbitrarily and arbitrarily applicable) and discuss how the latter may not only provide an explanation for stimulus equivalence, but for other types of derived stimulus relations as well.

Nonarbitrarily Applicable Relational Responding (NAARR)

Mammals, birds, fish, and insects can all be trained to respond to the relations between and among stimuli in the environment. However, for many different species, these relational responses appear to be characterized by two key properties: (a) they are rooted in a prior history of direct experience and (b) they are defined by the physical features of the to-be-related stimuli themselves (Giurfa, Zhang, Jenett, Menzel, & Srinivasan, 2001; Harmon, Strong, & Pasnak, 1982; Reese, 1968). RFT refers to this type of behavior as an instance of nonarbitrarily applicable relational responding (or NAARR) because the organism is relating stimuli based on their formal or physical properties. Properties such as color, shape, quantity, and size are considered ‘nonarbitrary” because they are based on the physical characteristics of the stimuli, unlike arbitrary or arbitrarily applicable properties, that are determined by social convention.

To illustrate the concept of NAARR more clearly, imagine that a pigeon is exposed to a learning task in which a sample stimulus (e.g., a red circle) is presented at the top of a computer screen and two comparison stimuli (e.g., a red and a green circle) are presented at the bottom of the screen. On trials where a red circle serves as a sample stimulus, selecting the red circle from the available comparisons is reinforced, and whenever a green circle is the sample, selecting the green circle from the available comparisons is reinforced. Training continues in this way across a wide spectrum of different colors and shapes. Once the bird is consistently correct across a large number of trials it is then presented with a number of entirely novel stimuli (that were never directly reinforced in the past). Research suggests that the pigeon will continue to select a shape from the bottom of the screen that is physically identical to a shape at the top of the screen even when that particular response was never previously reinforced (Frank & Wasserman, 2005). Now consider a series of studies wherein adult rhesus monkeys (Harmon et al., 1982) or marmosets (Yamazaki, Saiki, Inada, Iriki, Watanabe, in press) are trained to select the taller of two items that differ only in terms of their respective height. When subsequently presented with a previously “correct” item (i.e., a stimulus that was taller than its comparison stimulus) as well as a novel item that is even taller, they consistently select the latter, despite reinforcement for choosing the former at an earlier point in time. These studies, in addition to many others, suggest that animals can respond to the nonarbitrary (i.e., physical) relationship that exists between stimuli. In the above examples, pigeons related shapes based on their physical similarity to one another, while rhesus monkeys or marmosets responded to the comparative relationship between items that differed in their respective height (both providing examples of NAARR).

Arbitrarily Applicable Relational Responding (AARR)

RFT argues that while humans and nonhumans can both show NAARR, the former are typically exposed to a set of contingencies by the socio-verbal community that results in the development of a more advanced type of relating known as AARR. This behavior is not based on the physical relationship that exists between stimuli, but rather on the ability to derive relations between stimuli and events independently of their physical characteristics and in the absence of any direct training or instruction to do so.

As an example, imagine that I show you three identically sized coins and tell you that ‘*Coin A* is worth far less than *Coin B* which is, in turn, worth far less than *Coin C*.” I then given you the option to select any of the three coins and your hand immediately gravitates towards the third option. It is likely that the selection of Coin C occurs after you have derived a number of untrained relations between the various stimuli (e.g., that “Coin C is worth far more than A or B” and that “Coin A is worth far less than B or C”). What is remarkable here is that you can respond to the relation between stimuli despite the fact that (a) you have never encountered these items in the past, and (b) the three coins do not differ in any physical way. According to RFT, this example showcases an instance of AARR in which stimuli are arbitrarily related along a comparative dimension (worth). In a similar way the equivalence phenomenon discovered by Sidman can also be viewed as an instance of AARR, but one in which stimuli are arbitrarily related based on their sameness or similarity.

As we have seen above, the terms originally devised by Sidman to describe simple instances of AARR fail to accommodate the different properties of relational responses at increased levels of complexity. Thus a more generic set of terms was needed that could account for all possible derived relations that might be established between and among events. With this in mind, RFT argues that all derived stimulus relations (including equivalence) are characterized by three core properties known as (a) mutual entailment, (b) combinatorial entailment, and (c) the transformation of stimulus functions.

Mutual entailment. The first of these properties (*mutual entailment*), like the concept of symmetry, refers to the inherent bidirectionality or "reversibility" of stimulus relations, so that if A is related to B, people will also respond as if B is related to A. However, unlike symmetry, mutually entailed relations are not always symmetrical. If a person learns that the word for "*woman*" (A) is the same as the word for "*vrouw*" (B), they will likely respond as if the word '*vrouw*" (B) is the same as the word "*woman*" (A). Yet if they learn that a medicine in the blood (A) prevents cancer (B), they will not derive that cancer (B) prevents medicine in the blood (A). Thus while symmetry represents a subtype of mutual entailment in which derived and directly trained relations are the same, other mutually entailed relations are also possible that are not equivalent in nature.

Combinatorial entailment. The second property of derived stimulus relating is known as *combinatorial entailment*. This term refers to the fact that when two stimulus relations combine, a number of novel untrained relations tend to emerge. For instance, if A is related to B and B is related to C, then people will also respond as if A is related to C and C is related to A without any training or instructions to do so. Once again, it is important to note that while combinatorial entailment bears similarity to the concepts of ‘transitivity” and “equivalence,” it actually extends beyond both of them. To illustrate, imagine that you spend a summer in Europe and learn that the Croatian word for apple (A) is “jabuka,” (B) while in Spain it is “manzana” (C). In this instance, you will likely respond to a jabuka (B) as being the same as a manzana (C) and a manzana (C) as the same a jabuka (B). Now imagine that the following year you decide to travel to Canada where you are informed that a quarter (A) is worth more than a dime (B), and that a dime (B) is worth more than a nickel (C). In this instance, it is unlikely that you will treat these three coins as interchangeable or equivalent to one another. Rather a *bigger-than* relation will be derived between the quarter (A) and nickel (C) while a *smaller-than* relation will be derived between the nickel (C) and quarter (A). Therefore, while transitivity represents a subtype of combinatorial entailment based on similarity, other types of combinatorial relations are possible that are nonequivalent in nature.

Transformation of stimulus functions. The third and final property of derived stimulus relating, known as the *transformation of stimulus functions*, is particularly important from an RFT perspective because it is the process by which stimuli and events come to acquire, change, and lose their psychological properties. This term refers to the finding that when stimuli are related to one another - and the functions of one of those stimuli is modified in some way - the corresponding functions of other stimuli in that relation will spontaneously change without any training or instruction to do so. Critically, the transformation of functions always depend on the nature of the relation established between and among stimuli (e.g., Cahill et al., 2007; Gil et al., 2012; Smyth et al., 2006; Whelan, Barnes-Holmes & Dymond, 2006). As we pointed out above, when symmetry (*X-Same-Y*) and equivalence relations are formed (*X-Same-Y*; *Y-Same-Z*) and a function is then established for one of those stimuli (X), corresponding functions may subsequently *transfer* to the other stimuli in that relation as well (e.g., Y and Z). Imagine, for example, that a symmetry relation is formed between the word "poisonous" (X) and a novel liquid (Y) and a second relation is then established between novel liquid (Y) and a gas (Z). The formation of an equivalence relation between these three stimuli may lead to a transfer of functions from X to Y and Z, such that people respond with fear and attempt to avoid all contact with both substances, despite having never encountered either in the past. On the other hand, when nonequivalent relations are formed, the functions of a stimulus may not simply transfer but rather be *transformed* through those relations. Now imagine that an opposition relation is established between the word poisonous" (X) and a liquid (Y) and a second opposition relation is then established between the liquid (Y) and a gas (Z). Unlike before, the liquid (Y) will not evoke fear and avoidance (given that it is opposite to "poisonous"), while gas (Z) might (given that the combinatorially entailed relation is one of similarity between X and Z). In other words, while *transfers* of function constitute a subtype of *transformations* of functions (in which the psychological properties of a stimulus are broadly similar for stimuli in that relation), other types of derived stimulus relations can involve complex changes in functions (for the first empirical demonstration of a transformation of functions see Dymond and Barnes, 1995).[1]

Contextual Control Over AARR

If humans have the capacity to arbitrarily relate any stimulus to any other stimulus and

---

[1] An alternative term that captures the property of reflexivity does not appear to be necessary. Furthermore, some researchers have questioned the utility of reflexivity as a defining property of AARR because such responding may be based upon either derived stimulus relations or formal similarity (Steele & Hayes, 1991; see also Barnes, 1994). In any case, this issue is not important in the context of the current chapter and thus requires no further discussion.

substitute one for the other, then why does this ability not lead to complete and utter chaos? For example, why do people not try to eat the word “apple”, lick the words “ice cream” off a page, or even swat the word “fly” from a book? RFT proposes that the way in which people relate stimuli (and transform functions through those relations) is under the control of stimuli in the past or present environment known as contextual cues. While certain types of contextual cues specify how stimuli are related (e.g., “A *same as* B” or “A *causes* B”), others specify the psychological properties that are transformed through those relations (e.g., “A *tastes* disgusting” or “B *feels* soft”). RFT researchers usually refer to the former as “relational cues” (or C_{rels} for short) given that they specify how stimuli and events should be related. These cues can be used to relate stimuli in a near infinite number of ways, from relations based on similarity or opposition to those based on hierarchy, comparison, deictics, temporality, and/or causality. At the same time, responding can also be controlled by “functional cues” (or C_{funcs}) in the environment that specify the type of psychological properties that are transformed in accordance with stimulus relations. For example, the verbal stimulus “ice cream” could in principle evoke many of the psychological properties of actual ice cream (such as its taste, smell, appearance, or its coolness) based on the equivalence relation between the word and the food-item. If, however, someone asks you to *picture* ice cream, the visual properties of ice cream would likely predominate. Likewise, in the sentence “imagine what ice cream tastes like,” the expression “*tastes like*” may serve as a functional cue that is responsible for the fact that only the gustatory and not other functions of ice cream predominate (e.g., what it looks like).

RFT argues that all mutually and combinatorially entailed relations are under some form of contextual control, without which, different patterns of relational responding could not be observed. If AARR was not brought under the control of relational cues, for example, then all types of relations would apply to all events, resulting in chaotic and useless responses or “relational gridlock”. Likewise, if the transformation of functions were not restrained via functional cues, stimuli and events would collapse functionally in useless ways. For example, if all the functions of one stimulus in an equivalence relation were to transfer to another stimulus, then the two stimuli would merge and become indistinguishable in a psychological sense (e.g., a child would attempt to eat the word “candy” or an adult would try to drive the word “car”). Thus relational and functional cues may be seen as the metaphorical ‘scaffolding of relating’, specifying the manner in which stimuli should be related and functions transformed.

Summary

Relational frame theory is built upon a relatively simple idea: that a learned operant behavior known as arbitrarily applicable relational responding represents the basic functional “unit” from which phenomena like meaning, rule-governed behavior, and stimulus equivalence spring forth. The concept of AARR may appear to be difficult (and it is definitely technical), but most of its components have already been described. *Relational responding* refers to the ability to respond to relations between stimuli rather than just responding to each stimulus separately and RFT distinguishes between two classes of such behavior (NAARR and AARR). On the one hand, many different species can respond in novel ways based on the physical relationship that exists between previously encountered stimuli. In the language of RFT, such behaviors are defined as instances of *nonarbitrarily applicable* relational responding or NAARR. On the other hand, humans quickly learn how to (a) relate stimuli in ways that do not depend on their physical properties and (b) the stimuli involved in those relations often become related to each other in ways that were never explicitly trained. In the language of RFT, such outcomes are defined as instances of *arbitrarily applicable* relational responding or AARR. A behavior is defined as an instance of AARR whenever it shows evidence of mutual entailment, combinatorial entailment, and the transformation of function. RFT proposes that the manner in which stimuli are mutually and combinatorially related, as well as the psychological properties transformed in accordance with those relations, always depends on two different types of contextual control. The first (relational cues) specifies how stimuli are related while the second (functional cues) specify which functions are to be transformed through those relations.

Why is AARR so Important?

Over the past 40 years AARR has captured the imagination of behavioral scientists due to its symbolic, flexible, and generative properties. From the beginning researchers realized that this type of behavior was inherently *generative*. Providing humans with a small set of direct experiences consistently causes them to act as if those stimuli are related to one another in a staggering number of novel and untrained ways. Indeed, there is an exponential increase in the number of untrained relations as more and more stimuli are related, so that by the time eight stimulus relations are trained, people can - in principle - act as if those stimuli are related in several thousand untrained ways. Thus AARR represents a type of behavior that rapidly accelerates learning as more and more stimuli are related.

A second reason why AARR has attracted so much attention within the functional tradition is that it equips humans with an unparalleled degree of *flexibility* when interacting with the world around them. The aforementioned properties of relational responding allow organisms to better adjust to their environments, because relating itself becomes a part of the environment that increases the scope of the organism’s interactions with it. For instance, once an individual has learned how to respond in an arbitrarily applicable fashion, they can relate any stimulus to any other stimulus in a near infinite number of ways. They can relate stimuli with no physical resemblance (like spoken words, written words, and pictures) and these relations can come to control how they subsequently respond. People can also act as if stimuli have acquired, changed, or lost their psychological properties without the need to directly contact contingencies in the environment. To illustrate, suppose that a person learns that a novel item (A) is less than a second item (B) and that B is less than a third item (C). Thereafter, B is repeatedly paired with electrical shocks. Evidence indicates that people will display greater fear towards C than B and more fear to B than A, despite the fact that C and A were never paired with shock and that none of the stimuli share any physical similarity (Dougher, Hamilton, Fink, & Harrington, 2007). Moreover, if people learn that that they can avoid being shocked by repeatedly pressing a button when they see B, they will also press that same button when they see A and C (Auguston & Dougher, 1997). Finally, when avoidance of C is subsequently extinguished, participants will spontaneously stop avoiding A and B as well (Roche, Kanter, Brown, Dymond, & Fogarty, 2008; but see Luciano et al., 2013, 2014; Vervoort, Vervliet, Bennett, & Baeyens, 2014). Functionally speaking, it seems unlikely that this behavior is a simple case of stimulus generalization given that the three items bear no physical resemblance to one another. At the same time, it does not appear to be an instance of classical conditioning seeing as B and C were never paired with aversive events, nor an instance of operant conditioning given that fear or avoidance responding in the presence of certain stimuli was never reinforced in the past. Put another way, when organisms respond not only to external events but come to relate those events in different ways, the possibilities of manipulating and changing the world are dramatically increased.

In short, the generativity and flexibility of AARR, combined with its potential to scale in complexity, finally equipped researchers with a means to tackle psychological phenomena in way that was sorely lacking in the past (e.g., Skinner, 1957). Researchers quickly realized that two core features of AARR (generativity and flexibility) are also two core features of human language. For instance, the ability to derive relations between arbitrary stimuli closely mirrors the symbolic or referential nature of language, wherein spoken and written words share few physical properties with their referents, yet people respond to each of those stimuli as though they are equivalent (e.g., shouting “SNAKE” on an airplane might elicit many of the same fear responses as seeing a snake on an airplane). Likewise, the ability to derive a large number of relations between stimuli from a limited number of experiences also mirrors the remarkable generativity that lies at the heart of language (Hayes, Barnes-Holmes, & Roche, 2001). These theoretical observations were bolstered by empirical support on several fronts. Whereas verbally-able humans form derived stimulus relations with remarkable ease, their nonhuman counterparts have yet to demonstrate such relations convincingly or unequivocally (Lionello-DeNolf, 2009; Hughes & Barnes-Holmes, 2014). Individuals with verbal deficits also demonstrate impairments in their ability to respond in relationally complex ways (Barnes, McCullagh, & Keenan, 1990; O’Connor, Rafferty, Barnes-Holmes, & Barnes-Holmes, 2009) and providing remedial training in AARR can serve to address those deficits (Murphy, Barnes-Holmes, & Barnes-Holmes, 2005; Persicke, Tarbox, Ranick, & St. Clair, 2012; Walsh, Horgan, May, Dymond, & Whelan, 2014). At the same time, the ability to derive stimulus relations repeatedly correlates with cognitive and linguistic skills (Cassidy, Roche, & Hayes, 2011; O’Hora et al., 2008; O’Toole & Barnes-Holmes, 2009). The development of AARR initially emerges in infancy but develops gradually around the same time as verbal abilities (Luciano, Gomez-Becerra, & Rodriguez-Valverde, 2007) while brain-imaging studies reveal that derived relations produce similar patterns of neural activation to semantic processes (D. Barnes-Holmes et al., 2005; see also Whelan & Schlund, 2013).

As we shall see in the next chapter, the influence of AARR (and the history of learning that gave rise to it) is also argued to play a role in other domains such as perspective taking (McHugh & Stewart, 2012), implicit and explicit cognition (Hughes, Barnes-Holmes, & Vahey, 2012), problem-solving (Stewart, Barrett, McHugh, Barnes-Holmes, & O'Hora, 2013), analogical reasoning (Lipkens & Hayes, 2009), as well as fears, phobias, avoidance, and anxiety (Hayes, Strosahl, & Wilson, 1999). For now though let us turn to the origins of AARR and examine how this operant behavior may be learned early on in our development.

Part II: On the Origins and Properties of AARR

In the first section of the chapter we briefly discussed the research leading up to the development of RFT and provided an introduction to the theory itself. In Part II we take a closer look at the various assumptions that underpin this functional account of human language and cognition. We will show how the ability to frame events relationally is a type of operant behavior that is under both antecedent and consequential stimulus control. In effect, we argue that this operant is (a) generalized and purely functionally defined, (b) relational, and (c) arbitrarily applicable, but nonarbitrarily applied. In what follows we unpack each of these points in greater detail.

What is a Generalized and Functionally Defined Operant?

The concept of a generalized and functionally defined operant has often been used within the behavior analytic tradition to interpret or explain complex behaviors. When researchers speak of a functionally defined operant they are simply emphasizing the following point: that the core property of any operant (generalized or not) is the correspondence between a class of responses defined by its consequences and the variety of responses generated by these consequences. In other words, operant response classes are defined according to their functional effects as opposed to what any response within that class looks like (i.e., its topography). To illustrate this point, consider the act of powering on your computer. You may press the power button with your right finger, left hand, nose, a stick, or so forth. Although each of these responses appear different they are all button presses and qualify as members of the same operant class because they all share a common function (i.e., they all lead to the same consequence). The need to draw attention to the functional nature of operant classes arises from the fact that, in everyday life, the topographical and functional properties of operants often overlap, and it easy to confuse one with the other. In the above example, the operant of “powering on a computer” may be defined as the effect of activity on a certain button, but almost every such response involves the person using their right index finger. While that same button may also be activated in a variety of ways (e.g., by accidently dropping something on it) these responses are often ignored for practical purposes. And even if they were included, there is some notional limit to the range of topographies or movements that could possibly depress the button. The key point to appreciate here is that operant classes are defined functionally in terms of their effects rather than topographically based on what a given response looks like.

RFT puts this topography versus function issue squarely and unavoidably on the table (D. Barnes-Holmes & Barnes-Holmes, 2000; Healy, Barnes-Holmes & Smeets, 2000). It argues that in many cases the stimuli and responses that comprise an operant class have very few topographical features in common. For instance, it is possible to train people to emit entirely random sequences of numbers during an experiment by providing feedback on the randomness of a numerical string that participants had emitted on the previous trial. By definition, the functional class of “random number sequences” cannot be formed on the basis of what the stimuli look like because each of those stimuli vary in their topographical features. Yet this operant response class can still be trained (see Neuringer, 2002). While many other instances of generalized operants have now been identified (e.g., identity matching; Sidman, 2000; learning set; Harlow, 1959), the most well-known is arguably generalized imitation (Baer & Deguchi, 1985; Catania, 1998; Horne & Erjavec, 2007).

Generalized imitation refers to a specific functional relation: namely one between a model, an imitator, and a history of differential consequences for imitating. Although humans may be evolutionarily prepared to imitate the actions of other members of our species (Meltzoff, 2005), we seem to acquire a more general class of "do-what-other-people-do" as an operant behavior. Dinsmoor (1995) described the operative process as follows:

> When a number of correspondences have been reinforced between the actions of an observer and the actions of a model, the correspondence itself may become a governing factor in the relation between the two actions, extending to new topographies of behavior. (p. 264-265)

In other words, when the "correspondence" between the behavior of a model and the behavior of an imitator has been repeatedly reinforced, entirely novel responses may be imitated without the need for further reinforcement. Imagine, for example, that a father sets out to teach his child a range of behaviors (e.g., clapping, dancing, sharing) using a sock-puppet. If only one specific imitative response was ever trained (e.g., clapping) it is unlikely that generalized imitation would emerge, no matter how long the training lasted. However, if the relevant properties of the context are varied (e.g., the father engages in a range of different actions), consistent reinforcement is delivered for imitative responses, and increasingly novel and/or difficult responses are gradually introduced, then the functional class of generalized imitation will likely be acquired. A wide range of response topographies can now be substituted for the topographies used in the initial training, leading to a robust imitative repertoire. For instance, if a novel behavior is produced by the puppet (e.g., cleaning), the child may imitate this behavior despite the fact that this imitative response was never reinforced in the past. At this point the operant class is said to be generalized in that it contains imitative responses above and beyond those that were differentially reinforced. It is also functionally defined insofar as the stimuli and responses in that class bear no topographical similarity to one another – they are united by their common function.

AARR is a Generalized and Functionally Defined Operant

RFT argues that AARR is a generalized operant class that is established in a broadly similar way. Through early natural language interactions, human infants are exposed to a wide variety of stimuli, populations, and contexts in which differential consequences are provided for responding to the relationship that exists between stimuli. This functional relation is initially based on the nonarbitrary properties of the stimuli involved, but exposure to a sufficient number of exemplars of varying topography serves to abstract or “wash out” these irrelevant factors and brings the functional relation under the control of arbitrary contextual cues. These cues can be applied in such a way that stimuli can be related regardless of their physical relationship to one another. When different types of relating have been abstracted and brought under the control of contextual cues that extend beyond the physical properties of the related events, relational responding is said to be *arbitrarily applicable*.

To illustrate this point more clearly, imagine that you are attempting to teach your infant son how to name a number of objects around the house. You will likely begin by pointing at an item (e.g., a toy bear), uttering its name in the presence of your son, and then reinforcing any orientating response that he makes towards the item (i.e., hear the word *bear* → look at the bear). At the same time, you will also present the item to your son and then model or reinforce appropriate responses (see the bear → say the word *bear*). Both of these interactions will take place in the presence of contextual cues - and in natural language interactions - these cues typically take the form of questions such as "What is this?" or "What is the name of that?" In the language of RFT, you are directly reinforcing bidirectional responding in both directions to an object and its name in the presence of a contextual cue. Importantly, this training will not stop here: You and others in the social community (teachers, friends, family) will likely engage in the same exercise with your son across a vast spectrum of different objects, from toys ("*where is your bike?*"), to people ("*who is that?*"), food ("*this is an apple*"), and properties of the environment ("*that is called the sun…what is that called?*"), and do so in a wide variety of different contexts: at the park, home, at the shopping mall, school, and so on. Although the particular stimuli, people and contexts change across time, the functional relation between those stimuli is always held constant: Reinforcement is provided for relational responding in both directions and in the presence of arbitrary contextual cues. Gradually, after a sufficient number of exemplars, the generalized response pattern of object-word symmetry is abstracted from the topography of particular objects or events and comes under the control of contextual cues, thus establishing *derived* symmetry (i.e., being able to derive the untaught response when trained in only one direction) with any new word-object pair.

In other words, through a history of multiple exemplar training (MET), your son learns a type of generalized bidirectional responding that no longer depends on the physical features of the stimuli involved and that leads to the mutually entailed response being emitted in the absence of direct reinforcement. Now when you present him with a novel object and a vocal stimulus he has never encountered before (e.g., a laptop and the label "laptop"), he will respond in a bidirectional manner without any reinforcement for doing so (e.g., he will point to the laptop when asked "where is the laptop?" and answer with "laptop" when asked "what is this?"). According to RFT, this functional relation between an object and word constitutes an instance of mutual entailment in which stimuli are related on the basis of their arbitrary similarity to one another. In other words, your son has learned to treat a word and its referent as functionally similar in certain contexts.

A history of MET also allows for more complex relational responses to emerge. For instance, imagine a second scenario where you and your son examine a picture book containing many different items and you come across an entirely new stimulus: a picture of an African lion as well as the written word lion" printed on the opposite page. Given the prior history of reinforcement for bidirectional responding in the presence of contextual cues, pointing towards the picture and saying "*this is a lion*" will likely lead your son to emit a number of mutually entailed responses (e.g., asking "*what is that?*" will result in him saying "lion" while simply saying "*where is the lion?"* will lead him to point towards the appropriate picture). At the same time a second mutually entailed relation may be trained between the spoken and written words such that you utter the word "lion" in the presence of your son and then reinforce any orientating response that he makes towards the written word (i.e., hear the sound *lion* → look at the word lion). You will also orientate your son towards the written word (by pointing to it) and then modelling or reinforcing an appropriate response (see the word lion → emit the sound *lion*). Once again, these relational responses will be trained in all directions in the presence of certain contextual cues. For example, you will likely reinforce pointing towards the written word lion whenever your son sees a picture of a lion, pointing towards the picture whenever he sees the written word, and saying *lion* whenever he sees the written word or picture. In the language of RFT, your son is being exposed to a set of contingencies which reinforce bidirectional responding to the arbitrary relation between *two or more* stimuli. This same interaction will take place across a staggering number of different objects, words, and sounds in a variety of contexts. Although each of these relational responses may be reinforced initially, with sufficient training your son will come to emit the mutually and combinatorially entailed relations without any further training or instructions to do so. Thus, for example, when you relate a new picture of a zebra with the sound *zebra*, and the sound with the written word zebra, your son may respond to those stimuli as being related in a number of untrained ways (i.e., he will show evidence of mutual and combinatorial entailment).

The take home message here is that the ability to respond to the relation between stimuli can be discriminated, abstracted, and brought under arbitrary contextual control. In much the same way that training generalized imitation across multiple exemplars can lead to the abstraction of the functional relation between the model and observer, training humans to respond relationally across exemplars can lead to a situation in which relating itself (rather than the properties of the stimuli involved) becomes the important factor. In order for this to occur, the organism must be exposed to a sufficient number of exemplars that allows it to discriminate between the relevant features of the relation (responding to one event in terms of another based on a contextual cue) and the irrelevant features (the actual physical properties of the objects being related). As relational responding is freed through abstraction from the formal properties of related events, it comes under the control of relational and functional cues (C_{rels} and C_{funcs}) that serve as discriminative stimuli for the relevant relational response. When such cues are presented, the individual's prior history of relational learning can be brought to bear on any arbitrarily chosen set of stimuli, regardless of their nonarbitrary properties or the nonarbitrary relations between them. Moreover, while these bidirectional relations between stimuli are initially reinforced in both directions (e.g., “A is related to B” and “B is related to A”) the entailed or derived relations quickly come to be emitted without any further reinforcement for doing so (e.g., people will relate B to A whenever they learn that “A is related to B”). Thus, RFT suggests that the well-established concept of the operant can be extended to relational responding in order to explain one of the key generative features of human language. Indeed, from an RFT perspective, AARR is the behavioral process that underlies the symbolic nature of language and we will return to this point in Chapter Y.

For now it is worth noting that a history of differential reinforcement for bidirectional responding across multiple exemplars (MET) may also give rise to many other patterns of relational responding. Comparative relations provide a ready illustration. A parent might present a child with two boxes of toys, one with more toys than the other, and reinforce the selection of the box with more items in the presence of contextual cues such as “Which box has more toys?” or “Give daddy the box with more toys.” They may also reinforce the selection of the physically smaller object in the presence of cues such as “Which box has less toys?” or “Give daddy the box with less toys.” This training continues across many different exemplars that vary in their physical quantity and across many different contexts. However, in each case the functional relation of responding comparatively to the stimuli based on their physical properties is held constant. When the child begins to respond correctly to novel stimuli based on their physical size, relational training may shift to entirely arbitrary stimuli and continue until responding comes under the control of cues other than the physical dimensions of the stimuli involved. For instance, the parent may present their child with a nickel (which is physically larger than a penny), a penny and a dime (which is physically smaller than a nickel or a penny) and ask "which coin is worth the most?" The child may initially respond to the stimuli based on their nonarbitrary or physical properties and select the nickel because it is physically larger than either a penny or a dime. Such a response will likely fail to produce social reinforcement in that the parent may respond with "No -- the nickel is not worth the most." Given a sufficient number of trials, responding may thus be brought under the control of arbitrary contextual cues (e.g., the word *most*), such that the child now responds to the coins based on their conventional value rather than their physical size (i.e., the child selects the dime which is physically the smallest but monetarily the largest). The child may subsequently respond to foreign currencies in a functionally similar manner (i.e., by asking what individual coins are worth rather than assuming that larger coins are worth more than smaller coins). As we shall see in Part III, this type of training may also provide the basis for responding in accordance with distinction, opposition, hierarchy, spatial, and many other types of derived stimulus relations as well.

Summary

In short, RFT argues that AARR is a learned operant behavior that emerges via a protracted history of differential reinforcement across multiple exemplars and is characterized by three important properties: (a) it is generalized, (b) relational, and (c) arbitrarily applicable but nonarbitrarily applied. These three features do not undermine the argument that AARR is an operant behavior nor do they require that we invent a new type of operant to accommodate this phenomenon. However, they do require us to be conceptually precise in our understanding of what constitutes an operant (for a detailed discussion of this issue see D. Barnes-Holmes & Barnes-Holmes, 2000). AARR is *generalized* insofar as it is defined functionally in terms of its effects rather than the topography or form of any given stimulus or response. It is *relational* insofar as it is an operant that involves responding to one event in terms of another.

Part III: The Rich Complexity of AARR

In the previous section we argued that AARR represents a type of generalized operant behavior in which stimuli are related under the control of contextual cues that have themselves been abstracted through a history of differential reinforcement and brought to bear so that stimuli can be related to one another without regard to their physical properties. And as we outlined in Part I, once this ability to AARR has been acquired, people can relate stimuli and events in a near infinite number of ways, from relations based on sameness or coordination (e.g., "*Hond is the same as dog*") to those that involve comparison (e.g., "*Italian is better than French cuisine*"), opposition ("*night is opposite to day*"), temporality ("*summer comes before winter*"), hierarchy ("*sunflowers are a type of flower*"), analogy ("*I'm right as rain*"), and deictics ("*I am sitting here in this chair now*").

In the language of RFT, these different patterns of AARR are known as *relational frames*. This term is based on the metaphor of a "picture frame" and is used to convey the idea that people interact with the world by "framing events relationally." In much the same way that a picture-frame can hold a variety of images regardless of what those images actually look like (e.g., family photos, vacation images, or classical art), people can arbitrarily relate stimuli regardless of what they look, smell, feel, taste, or sound like. The key point to remember here is that "relational frames" are not hypothetical entities or mediating mental mechanisms used to account for behavior. Rather they are convenient labels for a specific type of AARR that: (a) shows the contextually controlled properties of mutual entailment, combinatorial entailment, and transformation of functions, (b) is due to a history of relational responding relevant to the contextual cues involved, and (c) is not based solely on a direct history of learning or the nonarbitrary characteristics of stimuli/responses. In other words, the terms *AARR* and *derived stimulus relating* are generic labels that are used to describe a type of generalized operant behavior while the terms *relational frames* or *relational framing* describe specific instances of that behavior (e.g., "stimuli were related in a frame of coordination, comparison, distinction…" and so forth). Over the past decade a wide variety of relational frames have been identified and subjected to an experimental analysis. In what follows, we shine a light on this work, discuss the main "families" of relational frames, and focus on the defining characteristics that distinguish one frame from another. While this treatment is not an exhaustive one (for a more detailed overview see Luciano, Rodrigquez, Manas, & Ruiz, 2009), it will serve to demonstrate some of the more common frames and how they may be combined to establish various classes of events.

Coordination

This relational frame is perhaps the most commonly known and ubiquitous pattern of relational responding and involves relating stimuli on the basis of identity, sameness, or similarity. Broadly speaking, stimuli within coordination relations are arbitrarily related under the control of cues such as "is" or some functional equivalent (e.g., "same as," "similar to," "like," "equals," or "means"). Thus, if an individual learns via experience or instruction that the English word "emergency" is the same as the French word "urgence," she will act as if "urgence" is the same as "emergency," even though this latter relationship has never been directly instructed. If she is then taught that "urgence" is the same as the German word "notfall," she will show evidence of mutual and combinatorial entailment (e.g., she will act as if "emergency" is the same as "notfall," "notfall" the same as "emergency" and "notfall" the same as "urgence"). Coordination appears to be the simplest arbitrarily applicable relational response and the one upon which many other relational frames are built.

Transformations of function in accordance with coordination relations are observed when a response trained in the presence of one stimulus also occurs in the presence of other stimuli that participate in that derived relation. Moreover, and unlike many other types of relational frames, the function acquired by each stimulus in the relation will be broadly similar (i.e., a transfer rather than a transformation will take place). Consider the above example in which a coordination relation was formed between the English word *emergency*, the French word *urgence* and the German word *notfall*. Once this relation has been established a person who visits France and hears loud shouts of “urgence” (or “notfall” during a vacation to Germany) may come to experience heighted arousal or fear. They may also use those same words in those same countries to attract help from others. Now imagine that a coordination relation is established between three novel stimuli (A, B, C) and a musical-mood induction technique is used to generate happy or sad affective states in the presence of B. People may report feeling happy or sad in the presence of A and C as well. Experimental evidence for the transfer of functions via coordination relations has now been obtained across a variety of populations and procedures (e.g., Y. Barnes-Holmes, Barnes-Holmes, Smeets, & Luciano, 2004; see also D. Barnes-Holmes, Barnes-Holmes, Smeets, Cullinan, & Leader, 2004; Cahill et al., 2007; Dymond et al., 2008; Dymond, Schlund, Roche, & Whelan, 2014; Gannon, Roche, Kanter, Forsyth, & Linehan, 2011; Gómez, López, Martín, Barnes-Holmes, & Barnes-Holmes, 2007; Munnelly, Martin, Dack, Zedginidze, & McHugh, in press; Rodríguez-Valverde et al., 2009).

Accumulating evidence also suggests that when other contextual cues are absent, people tend to relate stimuli in ways that involve lower levels of relational complexity. This can be seen on the matching-to-sample (MTS) task that typically yields evidence of equivalence responding until other relational cues are introduced that specify alternative relationships between stimuli and events (see also Hughes, De Houwer, & Barnes-Holmes, 2014). Thereafter more complex relations may emerge (e.g., Dougher et al., 2007; also see Hughes et al., 2012)[2].

Opposition

A second and more complex frame is that of opposition that involves arbitrarily relating stimuli under the control of cues such as “opposite,” or “completely different.” Whereas coordination involves the abstraction of a particular dimension along which stimuli may be equated (“*Sun is the same as Sol*”), the latter requires the abstraction of a dimension along which stimuli may be differentiated. That is, frames of opposition involve stimuli being related in ways that differ in direction (and to the same degree) from some reference point along a specified continuum. Along the physical dimension of temperature, for example, cool is the opposite of warm, and cold is the opposite of hot. Stimuli can also be framed in opposition along a variety of arbitrary dimensions as well (e.g., “odd is the opposite of even,” “work is the opposite of play,” and “easy is opposite to difficult”). Three points are worth noting here. First, the relevant dimension along which stimuli are related may or may not be specified in frames of opposition. If you are told that “cold is the opposite of hot” then the dimension of temperature is clearly implied and yet you can also relate A as the opposite of B without any such dimension being stipulated. Second, transformations of function through opposition relations lead to different outcomes at the mutual and combinatorial levels. While mutually entailed opposition relations involve opposition (“*Dog-Opposite-Cat*” entails that “*Cat-Opposite-Dog*”), combinatorially

---

[2] Note that a near infinite range of stimuli (including spoken or written words, symbols, sounds, etc.) may come to function as contextual cues controlling the arbitrary relating of stimuli and events. Although the most common examples will be highlighted throughout this section, it is important to appreciate that the coordination of these words with many other words and phrases generates an almost infinite array of substitute stimuli that will also control a given pattern of relational responding. For instance, it is possible to establish nonsense words, arbitrary shapes, sounds, and tastes as contextual cues within the laboratory that function in the same as words such as “is,” “opposite,” “more/less than,” “belongs to,” and so forth. As always, the importance of a stimulus ultimately lies in its function rather than a particular topography.

entailed opposition relations involve coordination ("*Cat-Opposite-Dog-Opposite-Tiger*" entails that "*Cat-Same-Tiger*"). One implication is that frames of opposition should only develop after frames of coordination have been successfully acquired (see Y. Barnes-Holmes, Barnes-Holmes, & Smeets, 2004 for evidence to this effect). Third, because these frames involve opposition at the mutually entailed level and coordination at the combinatorially entailed level, the transformations of functions that occur within these frames lead to stimuli acquiring different functions depending on their location within the relation. If a child is told, for example, that a type of candy (A) tastes disgusting and that candy (A) is opposite to candy (B) and candy (B) is opposite to candy (C) he may rapidly approach and consume candy (B) and yet avoid any contact with (A) or (C). Stated more precisely, mutually entailed opposition relations between an aversive (A) and neutral stimulus (B) may lead to the latter acquiring appetitive functions. A second mutually entailed opposition relation between (B) and another neutral stimulus (C) may lead to the latter acquiring aversive functions (see Dymond & Barnes, 1996; Dymond et al., 2008; Roche, Linehan, Ward, Dymond, & Rehfeldt, 2004; Whelan & Barnes-Holmes, 2004a, 2004b; Whelan, Cullinan, & O'Donovan, 2005).

Distinction

Similar to coordination and opposition, frames of distinction involve relating stimuli along (a) some physical or arbitrary dimension that is (b) under the control of cues such as "different," "dissimilar," and "is not the same" (e.g., "a star is different than a planet" or "freedom is not the same as justice"). Critically, however, these frames are characterized by a number of properties that distinguish them from their counterparts. One such property is their lack of specificity: Whereas all of the relations in frames of coordination and opposition are specified, this is not the case for those that comprise frames of distinction. If you are told that

“Toyota” is the same as “Honda,” and “Honda” is the same as “Nissan,” you can determine what the mutual and combinatorially entailed relations are between each of these stimuli. Likewise, if you learn that “good” is opposite to “evil,” which is in turn opposite to “honest,” you can also determine the derived relations between these stimuli as well. Yet if are told that “Google” is different to McDonalds,” and “McDonalds” is different to “Apple,” you cannot determine what the relation is between “Google” and “Apple” (they may be different or they may be the same). In other words, frames of distinction involve relating stimuli that differ in direction as well as degree along some continuum. Transformations of function through distinction relations may also demonstrate greater levels of variability than other frames given this lack of specificity. To illustrate, imagine that fear-eliciting properties are established for a Pokémon character in the laboratory by repeatedly pairing it with a shock. Thereafter a mutually entailed distinction relation is established between this character and a second Pokémon (B) and another such relation is established between Pokémon (B) and a third Pokémon (C). While participants may naturally come to fear Pokémon A, they could respond in a wide variety of ways towards (B) and (C) (e.g., these stimuli could elicit more or less fear than A and be rated as more or less negative, neutral, or positive than each other or A). Frames of distinction have attracted considerably less attention in the literature relative to their coordination, opposition, and comparative counterparts. While no study has established this pattern of relational responding where it was previously absent, a number of studies have examined this frame as it relates to clinical and cognitive phenomena (see Foody, Barnes-Holmes, Barnes-Holmes, & Luciano, 2013; Dixon & Zlomke, 2005; O’Toole & Barnes-Holmes, 2009; Roche & Barnes, 1996).

Comparison

The family of comparative frames involve responding to events in terms of a quantitative or qualitative relation along some specified dimension. This relating is usually under the control of contextual cues such as “heavier/lighter,” “better/worse,” “larger/smaller.” Many specific subtypes of comparative frames exist and each is defined (in part) by the dimensions along which the relation applies (size, attractiveness, speed, and so on). For example, if I say that “an elephant is bigger than a lion”: and “a lion is bigger than a mouse,” then the stimuli can be compared along the dimension of size, and you can derive that “an elephant is bigger than a mouse” and that “a mouse is smaller than an elephant.” However, I could also tell you that a “lion is faster than an elephant and an elephant is faster than a mouse,” in which case the same stimuli can be compared along the dimension of speed, and you can derive that “the lion is faster than the mouse and the mouse is slower than the lion.” In other words, comparative relations are specified at the mutual and combinatorially entailed levels. This specification increases when the dimension along which stimuli are being related is quantified. For instance, if I told you that “An elephant is three times the size of a lion and a lion is three times the size of a mouse,” you could derive that the elephant is exactly six times bigger than the mouse and that the mouse is six times smaller than the elephant.

When stimuli participate in comparative frames they may acquire similar or entirely different functions depending on how they are related. Imagine you learn that a certain stimulus (A) is less than (B), which is in turn less than (C), and that (A) signals that you are about to receive a monetary reward. You may experience more arousal when you see B and even greater arousal when you see C despite the fact that neither stimulus signaled reward in the past. In this case a transformation of function through a comparative relation has led to stimuli acquiring similar (eliciting) functions that vary in their respective magnitude. Now image that you are told that a new house is more valuable than an old shack and that a mansion is more valuable than a house. You may come to evaluate the shack negatively, the house neutrally, and the mansion positively. In this case, a transformation of function through comparative relations has led to stimuli acquiring different (evaluative) functions from one another. Comparative relations have been the subject of significant empirical scrutiny, both historically in the animal literature (e.g., transposition represents an instance of nonarbitrary comparative relating; Reese, 1968), and more recently with comparative framing in human infants and adults (Y. Barnes-Holmes, Barnes-Holmes, Smeets, Strand, & Friman, 2004; Berens & Hayes, 2007; Cassidy et al., 2011; Dougher et al., 2007; Munnelly, Freegard, & Dymond, 2013; Murphy & Barnes-Holmes, 2009; Vitale, Campbell, Barnes-Holmes & Barnes-Holmes, 2012; Whelan et al., 2006).

Spatial Relations

This family of frames involves the abstraction of a spatial dimension along which stimuli may be related and often comes under the control of cues such as “here/there, “in/out,” “front/back,” and so on (e.g., “Walter is in the lab,” “Bart is on his skateboard”). These frames share many similarities to comparative relations insofar as they involve responding to events in terms of their directional displacement along a specified (spatial) dimension. Moreover, they typically imply or specify how stimuli should be related with regard to a reference point and this characteristic makes them quite specific. For example, if you are told that “Arnold’s Gym” faces the back of “Rocky’s café,” you could order the fronts and backs of both premises in a linear sequence (back door of the gym, front door of the gym, back door of the cafe, front door of cafe). This is because front and back doors are relative to each premise, and knowing the orientation of the two buildings implies a number of additional relations between these stimuli. While spatial relations have been tangentially examined in the context of deictic framing (McHugh & Stewart, 2012; Weil, Hayes, & Capurro, 2011), empathy (Vilardaga, 2009) and clinical phenomena (Vilardaga, Estévez, Levin, & Hayes, 2012; Villatte, Monestès, McHugh, Baqué, & Loas, 2010), they have yet to be subjected to an experimental analysis in and of themselves or instantiated in organisms where previously absent or weak.

Temporal Relations

This family of frames also shares many similarities to comparative relations insofar as they involve responding to events in terms of their directional displacement along a specified (temporal) dimension. Such relations often come under the control of cues such as "before/after," "now/then," and "soon/later" (e.g., "*night comes before day*," "*cover your eyes now as the eclipse will start soon*"). Critically, these frames differ in important ways from those discussed above. According to RFT, the experience and construction of time differs for organisms with and without a history of AARR. For the latter time is simply change - the transition from "one totality to another in which the second totality now stands on, evolved from, or in some sense includes the first" (Hayes, 1992, p.112). From this perspective, organisms without the ability to AARR experience change in a unidirectional manner, from a now to a new now or from this to a new this, but never from a new this to an old this. To illustrate, consider a pigeon in an operant chamber whose behavior (key pecking) is reinforced in the presence of a green light. First, there was an observed green light, then a peck on a key, then food was eaten. Later there was an observed green light, then a peck on a key, and the food was eaten. Still later there was an observed green light, then a peck on a key, and the food was eaten. In this scenario the pigeon directly experiences a sequence of events or an orderly procession from one act to another. Thus for the bird (and other organisms without a history of AARR) time is:

> *the past as the future in the present*. Based on a history of change ("past")

> the animal is responding in the present to present events cuing change to other events. It is not the literal future that is part of the psychology of the animal – it is the past *as* the future. (Hayes, 1992, p.113)

In other words, the only future that such an organism knows is the past that it has experienced.

This no longer applies when organisms learn how to respond in an arbitrarily applicable fashion. Once this ability has been acquired people can temporally frame stimuli and events in ways that are independent of their prior experience (e.g., "*I'm going to heaven after I die*" or "*My life will be so much better after I kill myself*"). Such relations lead to time being framed as a bidirectional dimension along which events can be ordered and sequenced, so that consequences in the distant past ("*My grandfather and father both died from smoking before I was born*"), present ("*My exam takes place now*"), or far future ("*Eating healthy now will increase my likelihood of living to old age*") can exert an influence on how we behave. In other words, for an organism with the ability to engage in AARR, time is the *past as the constructed future in the present*. Based on a history of deriving temporal sequences among events ("past"), the organism is responding in the present by constructing a spatial relation between two or more stimuli. It is not the literal future that is part of the psychology of the organism – it is the past *as* the future, but in this case the future is constructed on the basis of their ability to AARR (Hayes, 1992, p.114). In other words, AARR influences the overarching experience of time so that the past can now be reconstructed and the future imagined, planned for, and contemplated whenever stimuli are framed as coming before or after, now or then, and sooner or later than one another.

Several properties of this type of framing are worth noting. Similar to coordination and opposition relations, temporal frames are typically specified in nature, so that knowing "*March* *comes before April*" and "*April comes before May*," allows you to derive that "*March comes before May*" and "*May comes after March*." However, transformations of function through temporal frames are often unlike those seen in their counterparts. In the case of comparative relations, transformation of functions usually involve a change in the physical properties of responses to the transformed stimuli (e.g., responding with greater fear when you see a spitting cobra than a wasp). Yet transformations of function through temporal relations usually result in the presence or absence of the response as a whole (e.g., people usually have dessert *after* rather than *before* dinner, or put on their clothes *after* rather than *during* a shower). To illustrate, imagine that two temporal relations were established in the laboratory (e.g., A1-*Before*-B1-*Before*-C1 and A2-*After*-B2-*After*-C2) and that C1 occasions an unpleasant electric shock whenever A1 is selected before B1 and B1 is selected before C1. Also image that C2 occasions a reward (e.g., money) whenever it is selected after B2 and B2 is selected after A2. Participants may come to respond with fear towards C1 (and evaluate C2 positively) only when the aforementioned sequences are emitted. In this case, the presence versus absence of fear or evaluative responding depends on how stimuli are temporally related to each other.

Unfortunately, evidence of temporal framing and the history of learning needed to establish it is currently scarce. This class of relations has received far less attention than other frames in the RFT literature, with existing work focused on their implications for intelligence and rule-following, as well as their experimental induction in adult populations (e.g., O'Hora et al. 2004; O'Hora, Peláez, & Barnes-Holmes, 2005; O'Hora et al., 2008; O'Toole & Barnes-Holmes, 2009). It remains to be seen how these frames are initially established in the natural environment, how the unidirectional experience of change comes to be abstracted, and the bidirectional dimension of time constructed. Given their potential role in suicide (Hayes, 1992) and delayed gratification (Mischel & Ayduk, 2004), for example, closer attention to temporal relations certainly seems warranted.

Deictics

Another family of frames is those that specify a relation between stimuli from the perspective of the speaker. Growing evidence suggests that these deictic frames are comprised of three main types of relations: (a) spatial (HERE-THERE), (b) temporal (NOW-THEN), and (c) interpersonal (I-YOU). Whereas coordination, distinction, and comparative relations emerge based on what people learn about stimuli that are physically similar, dissimilar, or quantitatively different along some dimension, deictics are somewhat different. They are not abstracted from a nonarbitrary or physical referent, but rather from the invariance of the speaker's perspective: Framing events deictically can only be achieved with regard to a specific perspective or point of view. Although people are exposed to a history of reinforcement for relating in different ways across a wide variety of stimuli, situations, and settings, it is the constant division between the speaker (who is always HERE and NOW) and the to-be-related stimuli (that are THERE and THEN) that provides the environmental consistency upon which deictic relations are abstracted and arbitrarily applied. For instance, during and throughout their early interactions with the socio-verbal community, children will learn to respond to and ask questions like the following: "*What are you doing here?*," "*What am I doing now?*," "*What will you do there?*," and so on. The physical environment in which such questions are asked and answered will differ from occasion to occasion, but the patterns of interpersonal (I-YOU), spatial (HERE-THERE), and temporal relations (NOW-THEN) will be applied consistently, and as the case with other relational frames, these patterns will be abstracted over time.

Deictic framing represents one of the most active areas in the RFT literature at present.

This class of relations has been found to emerge during early to middle childhood (McHugh, Barnes-Holmes, & Barnes-Holmes, 2004) and has been experimentally engineered where previously absent or weak (Rehfeldt, Dillen, Ziomek, & Kowalchuk, 2007; Weil et al., 2011). Deictics have also been implicated in a wide range of social and clinical phenomena, from social anhedonia (Villatte, Monestès, McHugh, Freixa i Baqué, & Loas, 2008) to schizophrenia, (Villatte et al., 2010), empathy and stigma (Vilardaga, 2009), Theory of Mind (Y. Barnes-Holmes, McHugh, & Barnes-Holmes, 2004), deception (McHugh, Barnes-Holmes, & Barnes-Holmes, 2007), false beliefs (McHugh, Barnes-Holmes, Barnes-Holmes, & Stewart, 2006), pathological altruism (Vilardaga & Hayes, 2012), intelligence (Gore, Barnes-Holmes, & Murphy, 2013), and the sense of self (for a detailed treatment of deictic framing see D. Barnes-Holmes, Hayes, & Dymond, 2001; McHugh & Stewart, 2012).

Hierarchy

This family of frames refers to the fact that different relations can be related to one another in a hierarchical fashion and typically comes under the control of contextual cues such as “is attribute/part/member of,” or “belongs to” (e.g., “Croissant is a type of pastry and pastries are a type of food”). This type of framing may be characterized by a number of properties. One such property is transitive class containment; that is, the relations between the members of a category are transitive. For instance, if C is a member of B, and B is a member of A, then C is a member of A (e.g., all Irish Setters are dogs, and all dogs are animals; therefore, all Irish Setters are animals; Slattery, Stewart, & O’Hora, 2011). These frames also involve asymmetrical relations that emerge between and among members or categories of the same hierarchy. For instance, if category A contains category B, then category B does not necessarily contain category A (e.g., “motor vehicles” contain “cars,” but “cars” do not contain all “motor vehicles”).

Like so many other families of frames it seems plausible that hierarchical framing is also based, in part, on an appropriate history of NAARR. For instance, a child might learn, in one context, to relate objects based on whether they are a physically part of other things (e.g., “*your toe is part of your foot*”) and in another, to relate objects based on whether they contain other things (e.g., “*the toy box contains your ball, your teddy and your building blocks*”). Given a sufficient number of exemplars across a variety of settings and situations, this behavior may come under the control of contextual cues that are abstracted and applied arbitrarily to stimuli regardless of their physical relationship to one another. For example, if you are told that A contains B and B contains C, then you can derive that A contains C and C is contained by A, without any specific information about the actual physical properties of the stimuli involved or how they are actually contained. A transformation of function may occur in accordance with this relational frame if you are told that C is a highly toxic substance, in that you might be more willing to pick up A rather than B because two containers afford more protection than just one.

It is important to recognize that hierarchical framing can involve increasingly complex interactions among frames. For instance, a more complete example of hierarchical framing, than the simple one provided above, might be as follows: container A contains two separate containers, B and X, and each of these containers contains two substances; B contains C and D, and X contains Y and Z. In this case, A contains all other elements within the network, but B contains only C and D and X contains only Y and Z. Thus a difference relation may be derived between B and X (because they are separate containers) but frames of coordination may be derived between C and D and between Y and Z, because each pair is housed within the same container (see Figure 2). Now imagine that I tell you that C and D are both inert substances but both Y and Z are highly toxic. You might be relatively willing to pick up container A and container B, but less willing to pick up container X.

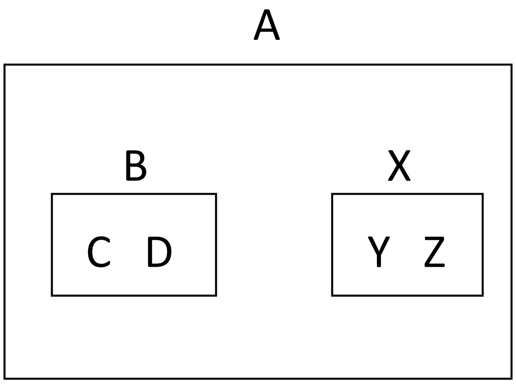

Figure 2. A graphical representation of one possible hierarchical relational network.

Hierarchical relational frames and the complex networks and transformation of functions that may emerge in accordance with them are ubiquitous in natural verbal behavior. Family or kinship relations provide a ready example. Imagine, for instance, that a friend informs you about a new television show about an American family known as the Simpsons. One part of the family is from Shelbyville and is aggressive, while the other part of the family is from Springfield and is funny. Thereafter, she tells you that Homer is from Springfield and Herbert is from Shelbyville, and upon hearing this, you may derive that Homer will be funny and Herbert aggressive. In the above example, the functions established for stimuli at one level of the hierarchy (i.e., the functions of aggressive and funny established for Springfield and Shelbyville) will alter the functions of subordinate class members (Homer is from Springfield and now funny, while Herbert is from Shelbyville and now aggressive) and superordinate class members (the Simpson's family is partly funny and partly aggressive).

The key point here is that a range of different relational frames can come to participate in hierarchical relational networks and this often leads to different transformations of function depending on how stimuli are related within those hierarchies. Take the previous example. When members of the Simpson clan are hierarchically framed according to their age, Homer may participate in a frame of coordination with his wife Marge, be comparatively higher than their children Bart, Maggie, and Lisa (who participate in a frame of coordination with each other), and comparatively lower than his father Abe. If you learn that old age is typically evaluated negatively then you may come to like Bart, Maggie, and Lisa more than Homer and Marge, and the latter stimuli more than Abe. At the same time, those stimuli could also be framed hierarchically with respect to gender, so that two superordinate classes of stimuli ("males" and "females"), which are part of the same hierarchy (gender), are framed in opposition to one another, while the members of those subordinate classes are framed on the basis of coordination (e.g., Bart is the same as Homer and Abe, who are all distinct from Marge, Lisa, and Maggie). If you then learn that girls are evaluated more positively than boys, an entirely different pattern of contextually controlled transformation of function will likely take place, with Marge, Lisa, and Maggie liked more than Homer, Bart, and Abe. These two hierarchical relations could also be combined so that stimuli are framed according to age and gender, leading to a more complex transformation of (evaluative) functions than before. The key point here is that complex hierarchical relational networks can involve different relations, and the manner in which those relations are hierarchically related will dictate how functions are transformed within and between those relations.

Similar to their spatial and temporal counterparts, hierarchical relations have received comparatively less attention than other families of relational frames (Gil, Luciano, Ruiz, & Valdivia-Salas, 2014; Gil et al., 2012; Griffee & Dougher, 2002; Slattery et al., 2011; Slattery &

Stewart, 2014), and have not been established in populations where such abilities are absent or weak. Nor has the importance of specific frames for the development of hierarchical relating been explored in the developmental literature (although for a discussion on how this might be achieved see Luciano et al., 2009). Given the importance of these frames for complex behaviors such as category learning (Murphy, 2002), evaluation (Hughes et al., 2014) and problem solving (Stewart et al., 2013), future work will need to pay closer attention to this pattern of relational responding.

Summary

In the previous section we highlighted a number of relational frames that have occupied the attention of RFT researchers over the past two decades. It seems important to repeat that these frames are not mediating mental or physical constructs, but simply labels that are used to talk about specific instances of a generalized operant behavior (AARR). Whenever researchers speak of a *relational frame* they are speaking of an organism, who given the proper historical and situational context, is relating stimuli independently of their physical properties and under the control of contextual cues that have previously been abstracted and are now being arbitrarily applied.

Research conducted over the last 20 years has refined our understanding of relational framing and the history of learning needed to produce these outcomes. This work has shown that stimuli can be related in many different ways, including coordination, opposition, distinction, comparison, deictics, temporality, and hierarchy. Indeed, a number of frames such as deictics, hierarchy, and temporality were only tentatively discussed by Hayes, Barnes-Holmes, and Roche (2001) in the original RFT book and yet these frames have now received empirical support as well as being implicated in a wide range of social and cognitive phenomena. Likewise, frames that were only ever examined with adult populations have now been established *ab initio* in the laboratory for organisms that previously lacked those abilities (infants and children). A common thread running through much of this work is that relational frames tend to emerge in-line with RFT’s predictions (i.e., via a systematic transition from nonarbitrary to arbitrary responding based on multiple-exemplar training with novel stimuli).

Despite these developments, a number of important issues still need to be addressed. Whereas coordination, comparison, opposition, and deictic relations have been subjected to tightly-controlled experimental analyses, their temporal, spatial, hierarchical, and causal counterparts have not enjoyed such attention. Only a handful of studies have examined these relations and often only tangentially in the context of other relational frames. Thus it still remains to be determined, for example, to what extent temporal and causal frames overlap functionally and whether it is better to consider them as largely separate or tightly connected. Likewise, we are only beginning to understand the manner and order in which different frames emerge in the natural environment (Luciano et al. 2009) as well as the role that specific frames play in maintaining and undermining other patterns of relational responding (Foody et al., 2013). It seems probable that relational repertoires like coordination and opposition need to be acquired prior to the emergence of comparison and more complex frames such as deictics, hierarchy, or causality. At present, many of these questions still await an answer. What has become clear, though, is that the aforementioned families of relational frames, while certainly important, are by no means the only ways in which humans can relate. Several researchers have started to model other types of frames such as those involved in mathematical relations (McGinty et al., 2012; Ninness et al., 2006, 2009). If RFT is correct, the number of relational frames is limited only by the creativity of the “relational” community that trains them.

Part IV: Evidence for AARR as a Learned Operant Behavior

So far we have argued the ability to respond in an arbitrarily applicable fashion unfolds through on-going interaction with the socio-verbal community (i.e., it is a type of behavior that can be generated, maintained, modified, or eliminated). If this assumption is correct, then AARR should demonstrate the same characteristics as any other operant: It should evolve gradually over time; be amenable to change and fall under the control of antecedent and consequential stimuli. In the following sections, we submit these theoretical claims to closer inspection and determine whether they hold up in the face of recent empirical evidence.

Development of AARR

Learning is an inherently developmental concept. As organisms interact with regularities in the environment, their actions gradually evolve and change (De Houwer, Barnes-Holmes, & Moors, 2013). If AARR is an instance of generalized operant behavior then it should also develop over time as well. At this point it is important to realize that any experiment that sets out to examine the development of AARR using adults or psychology students will involve relational performances that are almost certainly based on a prior history of relating. In such cases, the formation or modification of relational responses will be based on a rich and protracted history of relational learning. Consequently, these studies fail to provide strong evidence for the RFT view that AARR is established, in the first instance, as generalized operant behavior. Rather this requires that researchers shift their attention to the study of organisms with a limited or nonexistent history of AARR such as human infants and nonhuman animals. So far this analytic strategy has met with varying levels of success.

Development in humans. On the one hand, attempts to chart the emergence of AARR during the earliest stages of human development have proven fruitful. Studies indicate that the ability to engage in AARR is initially absent, but gradually grows in complexity, with mutually entailed coordination relations emerging first, followed by combinatorially entailed coordination, and noncoordination relations. For instance, several authors have found that infants are capable of receptive mutual entailment by 17 months, productive mutual entailment at 19 months, followed quickly by combinatorial entailment at 23 months (Lipkens, Hayes, & Hayes, 1993). Others have found that coordination relations can emerge even earlier when infants are provided with repeated training in symmetrical responding across multiple-exemplars (Luciano et al., 2007). Recent work with young children has sought to assess the development of AARR in a different way: by testing for the absence of certain patterns of relational responding and then establishing those very repertoires in the laboratory. Consider, for example, the work of Y. Barnes-Holmes, Barnes-Holmes, Smeets, Strand, & Friman, 2004. Prior to their studies a number of children were selected that had yet to learn how to respond in accordance with comparative or opposition relations. During the task they were shown a number of identically sized paper circles (referred to as "coins") and asked to pick the coin(s) that would buy as many sweets as possible. In Experiment 1 comparative relations were trained such that children were presented with three coins (A, B, C) and told: "*If this coin (experimenter points to the first coin -- A) buys less sweets than this coin (experimenter points to coin B), and this coin (experimenter points to B again) buys less sweets than this coin (Experimenter points to coin C), which coin would you choose to buy as many sweets as possible?*". In Experiment 2 opposition relations were trained. Children were once again shown three coins and told: "*If this coin (D) buys few sweets, and is opposite to this coin (C), and if this coin (C) is opposite to this coin (B), and if this coin (B) is opposite to this coin (A), which would you choose to buy as many sweets as possible?*"

After a protracted history of reinforcement for bidirectional responding across multiple exemplars, the children demonstrated evidence of derived opposition and comparative relating. In other words, they could relate any coin to any other coin in any direction, even when entirely novel coins and experimenters were introduced (see also Rehfeldt & Barnes-Holmes, 2009). This capacity to generate coordination, opposition, and comparative frames *ab initio* has now been replicated on numerous occasions (e.g., Y. Barnes-Holmes, Barnes-Holmes, & Smeets, 2004; Berens & Hayes, 2007; Gorham, Barnes-Holmes, Barnes-Holmes & Berens, 2009; Smeets & Barnes-Holmes, 2005; see also Barnes et al., 1990; Peláez, Gewirtz, Sanchez, & Mahabir, 2000).

Similar attempts to engineer deictic relations in young children have revealed that such relations are typically absent until 4 years of age (Y. Barnes-Holmes, McHugh, & Barnes-Holmes, 2004; McHugh, Barnes-Holmes, Barnes-Holmes, Stewart, & Dymond, 2007) and that their development can be accelerated via a similar history of learning as outlined above (Heagle & Rehfeldt, 2006), even in developmentally delayed populations who typically show deficits in this domain (Rehfeldt et al., 2007; Weil et al., 2011). As noted in the previous section, much of this work has focused on understanding perspective-taking, false belief, and deception as repertoires of derived relational responding and has employed cross-sectional developmental methodologies (see McHugh & Stewart, 2012). In each case, a clear developmental profile has emerged suggesting that the fluency of deictic relating is initially poor but quickly improves as a function of age. Comparable findings have also been obtained with tasks that require participants to relate derived relations to other derived relations (see Chapter Y for more details), which is evident in adults and 9 year old children, but absent or weak in 5 year old preschoolers (Carpentier, Smeets, & Barnes-Holmes, 2002; 2003; Pérez, García, & Gómez, 2011). When taken together, these findings support the notion that AARR is a generalized operant behavior that emerges through a protracted history of bidirectional training with multiple exemplars. This ability appears to be acquired in the earliest stages of infancy and rapidly scales in complexity, starting with mutual and combinatorial entailment based on coordination, and moving to comparative, opposition, deictic, and other patterns of relating during early to middle childhood. Providing normally-developing infants or young children (as well as their developmentally delayed counterparts) with MET appears to facilitate the development of relational abilities that were previously weak or absent.

Development in nonhumans. Researchers have also looked to nonhumans as another potential window into the emergence of AARR and much of this work has centered around the search for symmetry and equivalence responding in rats, pigeons, sea-lions, dogs, chimpanzees, bonobos, and baboons (for an overview see Lionello-DeNolf, 2009). While evidence has been obtained for symmetry and acquired equivalence responding in nonhumans, these performances can often be explained in ways that (a) do not involve AARR, (b) are only present in a subsection of the sample, or (c) are emitted with unacceptably low levels of accuracy (e.g., Dugdale & Lowe, 2000; Hayes, 1989a; Lionello-DeNolf, 2009). A number of authors have recently countered that these failures to observe symmetry and equivalence stem from properties of the procedures used and that nonhumans are in fact capable of such performances under a set of highly specific conditions (Galvão et al., 2005; Frank & Wasserman, 2005; McIlvane, Serna, Dube, & Stromer, 2000; Urcuioli, 2008; Zentall et al., 2014). It is worth noting that RFT has always remained agnostic to the possibility that AARR is a uniquely human capacity and it has never been argued that derived stimulus relating is forever beyond the grasp other species. Rather, RFT has simply viewed this claim as an empirical rather than purely theoretical one

(Dymond et al., 2003). What has emerged over the past 40 years is that so called associative symmetry and acquired-equivalence effects in nonhumans may arise, at least in part, from a functionally different history of learning to the generalized operant behavior of AARR displayed by their human counterparts (see Barnes & Roche, 1996; Hughes & Barnes-Holmes, 2014). Moreover, and as the evidence currently stands, it seems likely that there is also some “glass ceiling” in terms of relational complexity, contextual control, and generalizability that humans are capable of that is not evident elsewhere in the animal kingdom. For instance, other species have yet to show evidence for derived stimulus relating under nonequivalent contextual control and do not seem to respond to stimuli as being opposite, more than/less than, hierarchically, temporally, or casually related in an *arbitrarily applicable* fashion. Nor does it seem likely that they can relate derived relations to other derived relations and form increasingly complex networks of stimulus relations.

Nonetheless, investigations conducted with nonhumans under controlled conditions could provide a very useful platform for studying the early development of derived stimulus relating. Animal preparations and populations offer an opportunity to ask questions about AARR that cannot be answered with humans for ethical and practical reasons. This work could help us disentangle the history of learning involved in establishing and manipulating relational responding as a generalized operant behavior (e.g., Kastak & Shusterman, 2002; McIlvane, 2014). It could also provide information about the type, amount, and order of training that is required before relational responding becomes abstracted and generalizes to novel stimuli. Furthermore, it remains to be seen whether this advanced type of relational learning stretches across many different evolutionary branches or whether it is unique to a small number of species. The requirement for certain environmental or evolutionary conditions to be present before complex forms of AARR emerge also remains to be seen (Hayes & Long, 2013).

Flexibility of AARR

One of the hallmarks of operant behavior is its amenability to change. Accumulating evidence indicates that flexibility is a property of AARR as well. We now know that derived stimulus relations can be modified (even after they have formed) and that the relationships established between those stimuli may be altered individually or collectively depending on contextual factors. Imagine, for example, that a researcher establishes a coordination relation between three stimuli (A, B, C) and then modifies how all of those stimuli are related at a later point in time. A consistent finding is that novel relations will be derived that are in-line with these altered relations (e.g., O'Connor, Barnes-Holmes, & Barnes-Holmes, 2011). Yet if that same researcher only alters a small subset of the baseline relations then some of the derived relations will change, while others will remain intact (Cahill et al., 2007; Carr & Blackman, 2001; Dixon, Rehfeldt, Zlomke, & Robinson, 2006; Pilgrim & Galizio, 1995; Roche, Barnes, & Smeets, 1997; Watt, Keenan, Barnes, & Cairns, 1991; although see Garotti & De Rose, 2007; Pilgrim, Click, & Galizio, 2011; Smeets et al., 2003). These findings suggest that relating one event to another and combining relations among events are flexible behaviors under environmental control.

The fact that relating can come under different types of contextual control could also be seen as additional support for its flexibility. As we have seen, relating stimuli and events in the presence of contextual cues dramatically alters how people behave. After learning that the written word "poison" (A) is the same as a picture containing a unknown symbol (B) and that the latter is the same as the spoken word "G-I-F" (C), people will likely avoid consuming any items that contain images of B or that are labeled with C. However, if they subsequently learn that symbol (B) is the opposite of poison (A) and that GIF (C) is the opposite of (B) then they may approach the latter, but avoid substances labeled with the former stimulus. These relational responses are said to be flexible insofar as they vary systematically according to how stimuli are related, so that opposition relations (Dymond et al., 2008) can give rise to different outcomes to their comparative (Dougher et al., 2007), distinction (Foody et al., 2013), hierarchical (Gil et al., 2012), and temporal counterparts (O'Hora et al., 2004). In effect, derived stimulus relations themselves appear to be a type of relational flexibility.

AARR Falls Under Antecedent Stimulus Control

A third property of operant behavior is that it is falls under the control of the antecedent conditions that precede it. An extensive literature now indicates that AARR is also sensitive to antecedent stimulus control like any other operant. This is evident from the fact that a wide range of relational cues can be engineered in the laboratory and subsequently used to control how people relate stimuli and events to one another (as described in Part III).

AARR Falls Under Consequential Stimulus Control

The forth (and perhaps defining) property of an operant response class is that it is influenced by contingent consequences. Once again, AARR appears to fall under similar types of stimulus control. Consider the work of Wilson and Hayes (1996). In this study participants were exposed to a learning phase that was designed to establish three, four member coordination relations. They were then exposed to a second learning task that dismantled these relations and reorganized the stimuli involved into three new coordination relations. When the authors subsequently punished any response that was in-line with this second set of relations, a resurgence of the previously established coordination relations was observed (for similar findings see Healy et al., 2000). This resurgence effect is exactly the same as that seen elsewhere in the behavior analytic literature and refers to the fact that when an operant response ceases to produce reinforcement, or begins to produce punishment, responding becomes increasingly variable and earlier topographies tend to re-emerge (e.g., Doughty & Oken, 2008; Epstein, 1985; Lieving & Lattal, 2003). This work also highlights that "derived stimulus relations are extraordinarily difficult to break up, even with direct, contradictory training. …Once relations are derived, they never really seem to go away. You can add to them, but you cannot eliminate them altogether" (Pankey & Hayes, 2003, p. 315).

Further evidence that derived relational responding may be under consequential control was provided by Leonhard and Hayes (1991). In their study participants were first exposed to a MTS task designed to establish equivalence responding. They were then split into two different groups. The first group received a set of test trials that were entirely consistent with the previously formed equivalence relations, while the second group received a large number of trials that were inconsistent with those relations. Results revealed that these inconsistent trials significantly reduced mutual and combinatorial entailment in the second group. The authors also found that when all of the participants were subsequently trained and tested using novel stimuli and a normal MTS procedure, individuals from the second group continued to display significantly reduced levels of equivalence responding. These findings further support the argument that AARR is exquisitely sensitive to its consequences. Requiring people to relate stimuli in ways that are unspecified, inconsistent, or incoherent with their prior learning history serves to punish immediate and future instances of relational responding (see also Quinones & Hayes, 2014; Vitale et al., 2008).

At the same time a number of authors have found that the delivery of delayed consequences for responding can disrupt the emergence of mutual and combinatorially entailed relations (Healy, Barnes, & Smeets, 1998; Healy et al., 2000). For example, in Experiment 1 of the latter study, participants were exposed to a number of training and testing trials designed to establish two different coordination relations. Following this first cycle of training and testing they were provided with feedback about their performance on the task. While half of the participants received feedback that was consistent with the previously established relations (accurate), the other half received feedback that was inconsistent with those relations (inaccurate). Both groups were then exposed to another round of training and testing involving a novel set of stimuli. This cycle of training and testing, using novel sets of stimuli for each cycle, continued until a participant responded in-line with the feedback across three consecutive stimulus sets. Once this stability criterion was reached, feedback was switched from accurate to inaccurate (or vice-versa), and training and testing continued, using novel stimulus sets, until the performance once again reached the stability criterion. Results revealed that two contrasting patterns of AARR emerged as a function of the type of feedback (delayed consequences) provided. When feedback was accurate, participants responded in ways that were consistent with combinatorially entailed relations and when that feedback was inaccurate, they responded in ways that contradicted those same relations. In their final experiment, Healy and colleagues once again delivered accurate or inaccurate performance-contingent feedback. However, this time one type of feedback was provided following tests for mutual entailment and the other type of feedback following tests for combinatorial entailment. The authors found that derived stimulus relations, as behavioral units, could be “fractured” or “broken down” into component operants by appropriate reinforcement contingencies. In other words, they found that it is possible to separate and recombine mutually and combinatorially entailed relations by manipulating the type of feedback (consequences) that participants received – a finding that is broadly consistent with other studies (e.g., Gomez, Barnes-Holmes, & Luciano, 2001, 2002; Pilgrim & Galizio, 1995; Roche et al., 1997; Smeets et al, 2003). When taken together, these various lines of inquiry provide firm support for the notion that AARR is a generalized operant that is sensitive to consequential control.

Summary

Twenty years' worth of data in the areas of development, flexibility, antecedent and consequential control support RFT's claims that AARR is a form of learned operant behavior. Confirmatory evidence for this claim has now been obtained across a variety of normative and developmentally-delayed samples, stimulus modalities, settings, procedures, age groups, and relational frames. At the same time - and to the best of our knowledge - no contradictory evidence has been offered that seriously conflicts with the above account. What has become clear though is that derived relational responding represents a type of human performance that cannot be contained experimentally without raising serious ethical concerns. It develops so early and powerfully that researchers are often limited in their ability to test the processes that produce AARR in humans and to manipulate its emergence in systematic ways. Hopes that nonhuman animals would provide an alternative means for addressing these questions were seriously dampened by the finding that AARR is extremely difficult to observe elsewhere in the animal kingdom. Debate raged (and continues to do so) around the ability for other species to show even the most rudimentary properties of derived stimulus relating (mutual and combinatorial entailment) to the point that developmental research with infants and children (despite its methodological and ethical issues) has yielded greater insight into the operant nature of AARR.

Part IV: Important Additional Features of AARR

Since the first book length treatment of RFT by Hayes and colleagues (2001) a number of important features of AARR have increasingly attracted theoretical and empirical attention. These concern the influence of coherence in shaping how people frame events relationally, the complexity of the relational response itself, and the number of times that it has been derived in the past. In Chapter Y we will discuss the role that coherence, complexity, and derivation play in fast and slow cognition. But for now, let us examine these various features of AARR in greater detail.

Relational Coherence

As we have just seen, the ability to frame events relationally is a learned operant behavior that can be shaped, modified, or eliminated by contingencies of reinforcement or punishment. RFT argues that an important set of contingencies that serve to guide AARR in general are those in which social consequences are delivered to ensure that people frame in an internally consistent or coherent fashion. An individual is said to be responding *coherently* when all of the elements in a derived relation are related in a manner that is consistent with what was previously learned. In contrast, incoherent responding refers to instances in which derived relations are not consistent with what was previously learned. Imagine, for example, that a friend tells you “*an airplane is bigger than a car*’ and “*a car is bigger than a mouse*,” “*but a mouse is bigger than an airplane*.” You will quickly realize that stimuli are combinatorially entailed within this comparative relation in a manner that is incoherent with the mutually entailed relations and come to question the veracity of your friend’s statement. In other words,

> when one’s story does not cohere, the socio-verbal environment generally demands the story be changed until it is logically consistent. Stories that are

> consistent are generally reinforced (or set the stage for other reinforceable behaviors), while inconsistent stories typically result in no reinforcement or outright punishment. After lengthy exposure to these differential consequences, telling coherent stories increase in likelihood and incoherent storytelling becomes aversive. (Blackledge, Moran, & Ellis, 2009, p.243)

RFT proposes that from the cradle to the grave and in nearly every interaction in-between, the socio-verbal community reinforces coherent (and punishes incoherent) relational responding, to the extent that coherence itself quickly becomes a type of conditioned reinforcer for most individuals.

To date, only a handful of studies have examined the impact of relational coherence in the laboratory. Early work in this vein reported that people tend to revert back to previously learned, coherent ways of relating whenever they are faced with a situation in which they have to respond in an inconsistent fashion (e.g., Leonhard & Hayes, 1991; Pilgrim & Galizio, 1995; Wilson & Hayes, 1996). More recently, work has focused on how people react to relationally ambiguous contexts. For instance, learning that “Bill is more honest than George” and “Hillary is more honest than George” does not allow us to derive a coherent relation between Bill and Hillary: It may be that Bill is more honest than Hillary or vice-versa.

Quinones and Hayes (2014) recently sought to determine if people will respond in coherent or incoherent ways when faced with such ambiguous scenarios. In their first experiment, two comparative relations were established (A1 less than B1, A1 greater than C1; and A2 greater than B2, C2 less than A2) and participants were tested to determine if they would respond to B1 and C1 (as well as B2 and C2) as being the same/different and as being greater/less than one another. While the first relation is unambiguous (i.e., B1 is greater than C1 and thus B1 and C1 are different) the second relation is ambiguous (B2 and C2 may be the same or different). Consistent with past work, the authors found that people tend to derive coherent relations when presented with an unambiguous relation and yet responded idiosyncratically when presented with ambiguous relations (see also Vitale et al., 2008, 2012). That is, individuals who related B2 as different to C2 consistently related B2 as either greater or less than C2. In contrast, people who related B2 as the same as C2 responded randomly when they were asked to relate those same stimuli as either greater or less than each other. In a second experiment, participants received nonarbitrary multiple exemplar training that was designed to bias responding towards either "same" or "different" when confronted with an ambiguous stimulus relation. In line with Experiment 1, participants biased towards "different" produced consistent B2-C2 comparative responses, whereas those participants who were biased towards "same" responding produced idiosyncratic performances. When taken together, these findings suggest that an individual's learning history may bias them to respond in specific ways when confronted with ambiguous stimulus relations. Interestingly, even when participants responded idiosyncratically to the comparative relation, this pattern cohered with another response pattern (i.e., treating the B2 and C2 stimuli as the same). In this sense, a given instance of inconsistent responding may in fact be part of an overarching pattern of coherent relational responding, thus highlighting the potential power of coherence to function as a type of conditioned reinforcer for AARR itself.

Interestingly, while recent work has begun to unpack the relationship between coherence and AARR, evidence to support the claim that the former acts as a conditioned reinforcer for the latter remains extremely limited (Wray, Dougher, Hamilton, & Guinther, 2012). The fact that this question has yet to be subjected to a systematic analysis is somewhat surprising given that it is a popular assumption in the RFT literature and one that underpins many of the claims made about clinical (D. Barnes-Holmes, Barnes-Holmes, Cochrane, McHugh, & Stewart, 2004), cognitive (Hughes et al., 2012), and social phenomena (Roche, Barnes-Holmes, Barnes-Holmes, Stewart, & O'Hora, 2002). Future work in this area will need to articulate precisely when and how coherence functions as a reinforcer for AARR and examine its potential role in other psychological domains as well (Blackledge et al., 2009; Gawronski, 2012; Quinones & Hayes, 2014).

Relational Complexity

A second feature of AARR is the complexity of the relational response involved. As we have seen throughout this chapter, stimulus relations can vary in their complexity and be arranged along a continuum from low to high. Stimuli can be related to one another in a vast number of ways, from simple mutually entailed relations between single stimuli to combinatorial relations involving multiple stimuli, to the relating of stimulus relations to other relations, to the complex relating of entire relational networks to other networks. Not only can stimulus relations vary in their complexity, but so too can the type and number of functions transformed according to those relations. For example, mutually or combinatorially entailed relations between stimuli may involve single functions being transformed based on a relation between one stimulus and another, whereas the relating of complex networks of relations to other networks may involve a vast array of stimulus functions being modified in accordance with those relations.

Given that relational responses, like all behaviors, unfold across time, it appears that (all things being equal) more complex responses take additional time and are emitted with less accuracy relative to their less complex counterparts. To illustrate, consider the concept of nodal distance that refers to the number of nodes that link any two stimuli in a set of trained conditional relations. Interestingly, the time taken to respond in accordance with an equivalence relation increases and the accuracy of those responses decreases when the nodal distance within the equivalence class grows (Fields & Moss, 2007; Tomanari, Sidman, Rubio, & Dube, 2006; Wang, McHugh, & Whelan, 2012). Critically, however, when other relations above and beyond equivalence are involved, the complexity of a relation will be dictated not only by nodal distance, but by the number and type of relations involved (O'Hora, Roche, Barnes-Holmes, & Smeets, 2002; Steele & Hayes, 1991). Indeed, some work now indicates that as the number and type of relations increase, the speed and accuracy of responding decreases relative to responses that are at lower levels of complexity (see D. Barnes-Holmes et al., 2005; Hyland, Smyth, O'Hora, & Leslie, 2014; Reilly, Whelan, & Barnes-Holmes, 2005; Vitale et al., 2008).

Levels of Derivation

Relations can not only vary in their complexity, but also in the degree to which they have been previously derived in the past. As noted above, derivation refers to the finding that once a set of relations between stimuli is directly trained, a number of additional untrained relations also emerge and allow for the transformation of functions. To illustrate, consider a situation where a participant has just been trained to select B when given A and C when given B. Thereafter, and upon testing, a series of untrained relations are evident (e.g., selecting A when given C or C when given A). In this learning situation, the first instance in which the person derives the relation between A and C may be defined as a ‘‘high derivation’’ response given that the history of deriving that particular response is minimal. Alternatively, imagine that the same person is then provided with an ever increasing number of opportunities to derive the relation between those same stimuli. Across each of these successive derivation opportunities the resulting response may come to be increasingly defined as involving ‘‘low’’ levels of derivation. Note that, according to RFT, derivation may well decline, with repeated instances, even when some form of programmed reinforcement is not provided for each derived response because derivation itself is ‘‘rewarded’’ by contacting increased relational coherence (see Hayes, Fox, et al., 2001, pp. 42–43).

It also appears that the extent to which a response has been derived in the past will influence its probability of being emitted quickly and accurately in the future. For instance, the speed with which participants derive coordination, comparative, and opposition relations becomes significantly faster with each successive opportunity to derive (O’Hora et al., 2002; Roche et al., 2004; Steele & Hayes, 1991). Likewise, an overarching history of derivation may facilitate the emergence of more accurate relational performances within and across stimulus sets (Healy et al., 2000; Roche et al., 2004; Saunders & Green, 1999; Sidman, 1994; Wang et al., 2012; Wulfert & Hayes, 1988).

It should be noted here that the concept of “levels of derivation” in RFT may be applied to multiple levels of analysis. Imagine, for example, that a young child is trained to relate A *same as* B and B *same as* C and is then tested for the C *same as* A combinatorially entailed relation. The first time the child produces this relational response the level of derivation would be defined as high, but if the child derives that response many times thereafter, derivation is seen as dropping to lower and lower levels. Now imagine that the child is trained and tested for the same relational frame, but using a new set of stimuli (train D *same as* E and E *same as* F; test F *same as* D). Once again, the first time the child derives the F-D relation, derivation for that particular response would be defined as high, and would then be seen as dropping with each successive F-D relational response. Critically, however, at the level of the relational frame itself (in this case coordination), derivation would also be defined as dropping from responding with the A-B-C stimulus set to the D-E-F set. In other words, the level of deriving the frame of coordination itself may be seen as reducing across multiple stimulus sets. The same general logic applies to more complex patterns of AARR. Thus, for example, the first time a child derives a relation of coordination between two frames of coordination (C *same as* A is the same as F *same as* D) derivation will be high for both the particular relating-relations response and the act of relating derived relations itself. If relating-relations is then “practiced” across other novel sets of stimuli, the level of derivation involved in relating-relations would be seen as reducing across those sets. This view of AARR helps to make sense of the fact that complex relational responding, such as relating-relations, appears to be relatively weak in children aged 4 - 5 but thereafter appears to grow in strength as they are provided with more and more opportunities to derive such complex relational responses (see Stewart & Barnes-Holmes, 2004).

As noted in the introduction to this section of the current handbook, the recent focus on relational coherence, complexity, and levels of derivation in RFT is serving to inject a much needed emphasis on the role of reinforcement contingencies in understanding AARR. Indeed, we believe that this refocusing will be quite transformative in terms of moving RFT forward over the coming years, in that it will encourage researchers to identify the functional units of analysis that are being selected and strengthened or weakened as individuals interact with their verbal communities and the world around them. In time, we hope that the need for middle-level terms, such as defusion, acceptance, and even psychological flexibility itself, that currently abound in the applied wing of CBS, may be replaced with (or perhaps better supported by) RFT concepts that are far more closely tied to (experimental) functional analyses of the behavioral units that are actually selected by manipulable environmental variables.

Summary

On the one hand, it appears that humans search for and create consistency between and among derived stimulus relations involving arbitrary stimuli. Once the ability to engage in AARR emerges "it is maintained by *coherence* … when relational networks are internally coherent, we feel confident that we understand. Because such understanding often predicts an ability to control events, coherence becomes a proxy variable for instrumental success" (Hayes, 2002, p.104). On the other hand, it seems that the *complexity* of a relational response and the degree to which it has been previously *derived* can vary along a continuum from low to high. Like the concept of the relational frame discussed above, complexity and derivation are not hypothetical constructs or mental mechanisms: They are simply properties of AARR that will be more or less evident in different contexts. Specifically, it appears that the complexity of a relational response, as well as the degree to which it has been previously derived in the past, influences the probability that it will be emitted with speed and accuracy in the future. Responses characterized by an extensive history of derivation and low levels of complexity appear to be emitted with relatively greater speed/accuracy than their more complex and less derived counterparts (see Hughes et al., 2012 for a detailed treatment). Coherence, complexity, and derivation seem to play an important role in many areas of psychological science, an idea upon which we will shall expand in Chapter Y.

Conclusion

Throughout this chapter we sought to provide an accessible introduction to, and a state-of-the-art report on, RFT and the empirical work that it has stimulated over the past two decades. At the core of this account lies a relatively simple claim with far-reaching consequences - namely - that a generalized operant behavior known as arbitrarily applicable relational responding is learned early on in our development and provides the behavioral foundation for human language and cognition. So far our story has focused on the background, origins, and nature of AARR and largely left its role in human language and cognition untouched. We adopted this strategy so that the various components of AARR could be carefully considered and the empirical basis for this account examined before we demonstrate how it has been interfaced with specific aspects of psychological science. Let us now turn our attention, in the next chapter, to how this ability to frame events relationally provides the basis upon which many complex human behaviors are built.

References

Anderson, N. H. (2013). Unified psychology based on three laws of information integration. *Review of General Psychology, 17*, 125-132.

Aristotle. (1941). De Memoria et Reminiscentia, (J. I. Beare, trans). In R. McKeone (Ed.), *The basic works of Aristotle* (pp. 607-617). New York, NY: Random House.

Auguston, E. M., & Dougher, M. J. (1997). The transfer of avoidance evoking functions through stimulus equivalence classes. *Journal of Behavioral Therapy and Experimental Psychiatry, 28,* 181-191.

Baer, D. M., & Deguchi, H. (1985). Generalized imitation from a radical-behavioral viewpoint. In S. Reiss & R. R. Bootzin (Eds.), *Theoretical issues in behavior therapy* (pp. 179 –217). New York, NY: Academic Press.

Barnes, D. (1994). Stimulus equivalence and relational frame theory. *The Psychological Record, 44,* 91-124.

Barnes, D., McCullagh, P., & Keenan, M. (1990). Equivalence class formation in non-hearing impaired children and hearing impaired children. *Analysis of Verbal Behavior, 8,* 1-11.

Barnes, D., & Roche, B. (1996). Relational frame theory and stimulus equivalence are fundamentally different: A reply to Saunders' commentary. *The Psychological Record, 46,* 489-507.

Barnes-Holmes, D., & Barnes-Holmes, Y. (2000). Explaining complex behavior: Two perspectives on the concept of generalized operant classes. *The Psychological Record, 50,* 251-265.

Barnes-Holmes, D., Barnes-Holmes, Y., Smeets, P. M., Cullinan, V., & Leader, G. (2004). Relational frame theory and stimulus equivalence: Conceptual and procedural issues. *International Journal of Psychology and Psychological Therapy, 4,* 181-214.

Barnes-Holmes, D., Hayes, S. C., & Dymond, S. (2001) Self and self-directed rules. In S. C. Hayes, D. Barnes-Holmes, & B. Roche (Eds.), *Relational frame theory: A post-Skinnerian account of human language and cognition* (pp. 119-139). New York, NY: Kluwer Academic/Plenum.

Barnes-Holmes, D., Luciano, C., & Barnes-Holmes, Y. (2004a). Relational frame theory: Definitions, controversies, and applications I. *International Journal of Psychology and Psychological Therapy, 4,* 177-394.

Barnes-Holmes, D., Luciano, C., & Barnes-Holmes, Y. (2004b). Relational frame theory: Definitions, controversies, and applications II. *International Journal of Psychology and Psychological Therapy, 4,* 443-622.

Barnes-Holmes, D., & Murphy, C. (2007). Addressing the generativity of language: A late reply to Chomsky. In B. S. Mesmere (Ed.), *New autism research developments* (pp. 175–196). New York, NY: Nova Science Publishers.

Barnes-Holmes, D., O'Hora, D., Roche, B., Hayes, S. C., Bissett, R. T., & Lyddy, D. (2001). Understanding and verbal regulation. In S. C. Hayes, D. Barnes-Holmes, & B. Roche (Eds.), *Relational frame theory: A post-Skinnerian account of human language and cognition* (pp. 103-117). New York, NY: Kluwer Academic/Plenum.

Barnes-Holmes, D., Staunton, C., Whelan, R., Barnes-Holmes, Y., Commins, S., Walsh, D., . . . Dymond, S. (2005). Derived stimulus relations, semantic priming, and event-related potentials: Testing a behavioral theory of semantic networks. *Journal of the Experimental Analysis of Behavior, 84),* 417-430.

Barnes-Holmes, Y., Barnes-Holmes, D., McHugh, L., & Hayes, S. (2004). Teaching derived relational responding with young children. *Journal of Early and Intensive Behavior Intervention, 1,* 56-90.

Barnes-Holmes, Y., Barnes-Holmes, D., & Smeets, P. M. (2004). Establishing relational responding in accordance with opposite as generalized operant behavior in young children. *International Journal of Psychology and Psychological Therapy, 4,* 559-586.

Barnes-Holmes, Y., Barnes-Holmes, D., Smeets, P. M., & Luciano, C. (2004). The derived transfer of mood functions through equivalence relations. *The Psychological Record, 54,* 95-114.

Barnes-Holmes, Y., Barnes-Holmes, D., Smeets, P. M., Strand, P., & Friman, P. (2004). Establishing relational responding in accordance with more-than and less-than as generalized operant behavior in young children. *International Journal of Psychology and Psychological Therapy, 4,* 531-558.

Barnes-Holmes, Y., McHugh, L., & Barnes-Holmes, D. (2004). Perspective-taking and Theory of Mind: A relational frame account. *The Behavior Analyst Today, 5,* 15-25.

Baron, A., & Galizio, M. (1983). Instructional control of human operant behavior. *The Psychological Record, 33,* 495-520.

Bentall, R. P., Lowe, C. F., & Beasty, A. (1985). The role of verbal behavior in human learning: II. Developmental differences. *Journal of the Experimental Analysis of Behavior, 43,* 165–81.

Berens, N. M., & Hayes, S. C. (2007). Arbitrarily applicable comparative relations: Experimental Evidence for relational operants. *Journal of Applied Behavior Analysis, 40,* 45-71.

Blackledge, J. T., Moran, D. J., & Ellis, A. E. (2009). Bridging the divide: Linking basic science to applied psychotherapeutic interventions - A relational frame theory account of cognitive disputation in rational emotive behavior therapy. *Journal of Rational -Emotive & Cognitive-Behavior Therapy, 27,* 232-248.

Brino, A. L., Campos, R. S., Galvão, O. F., & McIlvane, W. J. (in press). Blank-comparison matching-to-sample reveals a false positive symmetry test in a capuchin monkey. *Psychology & Neuroscience*.

Cahill, J., Barnes-Holmes, Y., Barnes-Holmes, D., Rodríguez-Valverde, M., Luciano, C., & Smeets, P. M. (2007). The derived transfer and reversal of mood functions through equivalence relations II. *The Psychological Record, 57,* 373-389.

Carpentier, F., Smeets, P. M., & Barnes-Holmes, D. (2002). Matching functionally-same relations: Implications for equivalence-equivalence as a model for analogical reasoning. *The Psychological Record, 52,* 351-312.

Carpentier, F., Smeets, P. M., & Barnes-Holmes, D. (2003). Equivalence-equivalence as a model of analogy: Further analyses. *The Psychological Record, 53,* 349-372.

Carr, D., & Blackman, D. E. (2001). Relations among equivalence, naming, and conflicting baseline control. *Journal of the Experimental Analysis of Behavior, 65,* 245–247.

Cassidy, S., Roche, B., & Hayes, S. C. (2011). A relational frame training intervention to raise intelligence quotients: A pilot study. *The Psychological Record, 61,* 173–198.

Catania, A. C. (1998). *Learning* (4th ed.). Upper Saddle River, NJ: Prentice Hall.

Chomsky, N. (1959). A review of B.F. Skinner's *Verbal Behavior*. *Language, 35,* 26-58.

De Houwer, J., Barnes-Holmes, D., & Moors, A. (2013). What is learning? On the nature and merits of a functional definition of learning. *Psychonomic Bulletin & Review, 20,* 631 -642.

Devany, J. M., Hayes, S. C., & Nelson, R. O. (1986). Equivalence class formation in language-able and language-disabled children. *Journal of the Experimental Analysis of Behavior, 46,* 243-257.

Dinsmoor, J. A. (1995). Stimulus control: II. *The Behavior Analyst, 18,* 253-269.

Dixon, M. R., Rehfeldt, R. A., Zlomke, K. M., & Robinson, A. (2006). Exploring the development and dismantling of equivalence classes involving terrorist stimuli. *The Psychological Record, 56,* 83–103.

Dixon, M. R., & Zlomke, K. M. (2005). Using the precursor to the relational evaluation procedure (PREP) to establish the relational frames of sameness, opposition, and distinction. *Revista Latinoamericana de Psicología, 37(2),* 305-316.

Dougher, M. J., Hamilton, D., Fink, B., & Harrington, J. (2007). Transformation of the discriminative and eliciting functions of generalized relational stimuli. *Journal of the Experimental Analysis of Behavior, 88,* 179-197.

Dougher, M., Perkins, D. R., Greenway, D., Koons, A., & Chiasson, C. (2002). Contextual control of equivalence-based transformation of functions. *Journal of the Experimental Analysis of Behavior, 78,* 63-94.

Doughty, A. H., & Oken, G. Extinction-induced response resurgence: A selective review. *The Behavior Analyst Today, 9,* 27–34.

Dugdale, N., & Lowe, C. F. (2000). Testing for symmetry in the conditional discriminations of language trained chimpanzees. *Journal of the Experimental Analysis of Behavior, 73,* 5-22.

Dymond, S., & Alonso-Álvarez, B. (2010). The selective impact of Skinner's *Verbal Behavior* on empirical research: A reply to Schlinger (2008). *The Psychological Record, 60,* 355-360.

Dymond, S., & Barnes, D. (1995). A transformation of self-discrimination response functions in accordance with 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.

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

Dymond, S., & Roche, B. (Eds.), (2013). *Advances in Relational Frame Theory: Research and application*. Oakland, CA: New Harbinger.

Dymond, S., Roche, B., & Barnes-Holmes, D. (2003). The continuity strategy, human behavior, and behavior analysis. *The Psychological Record*, *53,* 333-347.

Dymond, S., Roche, B., Forsyth, J. P., Whelan, R., & Rhoden, J. (2008). Derived avoidance learning: Transformation of avoidance response functions in accordance with same and opposite relational frames. *The Psychological Record*, *58,* 269-286.

Dymond, S., Schlund, M. W., Roche, B., & Whelan, R. (2014). The spread of fear: Symbolic generalization mediates graded threat-avoidance in specific phobia, *The Quarterly Journal of Experimental Psychology, 67*, 247-259.

Epstein, R. (1985). Extinction-induced resurgence: Preliminary investigations and possible applications. *The Psychological Record*, *35,* 143–153.

Fields, L., & Moss, P. (2007). Stimulus relatedness in equivalence classes: Interaction of nodality and contingency. *European Journal of Behavior Analysis*, *8,* 141–159.

Foody, M., Barnes-Holmes, Y., Barnes-Holmes, D., & Luciano, C. (2013). An empirical investigation of hierarchical versus distinction relations in a self-based ACT exercise. *International Journal of Psychology & Psychological Therapy, 13,* 373-388.

Frank, A. J., & Wasserman, E. A. (2005). Associative symmetry in the pigeon after successive matching-to-sample training. *Journal of the Experimental Analysis of Behavior, 84,* 147-165.

Galvão, O.F., Da Silva Barros, R., Dos Santos, J. R., De Faria Brino, A. L., Brandao, S., Lavratti, C. M., Dube, W.V., & McIlvane, W. J. (2005). Extent and limits of the matching concept in Cebus apella: A matter of experimental control?.*The Psychological Record*, *55,* 219-232.

Gannon, S., Roche, B., Kanter, J., Forsyth, J. P., & Linehan, C. (2011). A derived relations analysis of approach-avoidance conflict: Implications for the behavioral analysis of human anxiety. *The Psychological Record, 61,* 227-252.

Garcia-Marques, L., & Ferreira, M. (2011). Friends and foes of theory construction in psychological science: Vague dichotomies, unified theories of cognition, and the new experimentalism. *Perspectives on Psychological Science, 6,* 192-201.

Garotti, M., & De Rose, J. C. (2007). Reorganization of equivalence classes: Evidence for contextual control by baseline reviews before probes. *The Psychological Record, 57,* 87 –102.

Gawronski, B. (2012). Back to the future of dissonance theory: Cognitive consistency as a core motive. *Social Cognition, 30,* 652-668.

Gil, E., Luciano, C., Ruiz, F.J., & Valdivia-Salas, V. (2012). A preliminary demonstration of transformation of functions through hierarchical relations. *International Journal of Psychology and Psychological Therapy, 12,* 1-19.

Gil, E., Luciano, C., Ruiz, F. J., & Valdivia-Salas, S. (2014). A further Experimental Step in the Analysis of Hierarchical Responding. *International Journal of Psychology and Psychological Therapy, 14,* 137-153.

Giurfa, M., Zhang, S., Jenett, A., Menzel, R., & Srinivasan, M. V. (2001). The concepts of 'sameness' and 'difference' in an insect. *Nature, 410,* 930-933.

Gomez, S., Barnes-Holmes, D., & Luciano, M. C. (2001). Generalized break equivalence I. *The Psychological Record, 51,* 131-150.

Gomez, S., Barnes-Holmes, D., & Luciano, M. C. (2002). Generalized break equivalence II: Contextual control over a generalized pattern of stimulus relations. *The Psychological Record, 52,* 203-220.

Gómez, S., López, F., Martín, C. B., Barnes-Holmes, Y., & Barnes-Holmes, D. (2007). Exemplar training and a derived transformation of functions in accordance with symmetry and equivalence. *The Psychological Record, 57,* 273-294.

Gore, N., Barnes-Holmes, Y., & Murphy, G. (2013). The relationship between intellectual functioning and relational perspective-taking. *International Journal of Psychology and Psychological Therapy, 10,* 1-17.

Gorham, M., Barnes-Holmes, Y., Barnes-Holmes, D., & Berens, N. (2009). Derived comparative and transitive relations in young children with and without autism. *The Psychological Record, 59,* 221-246.

Greer, R. D. (2008). The ontogenetic selection of verbal capabilities: Contributions of Skinner's verbal behavior theory to a more comprehensive understanding of language. *International Journal of Psychology and Psychological Therapy, 8,* 363-386.

Griffee, K., & Dougher, M. J. (2002). Contextual control of stimulus generalization and stimulus equivalence in hierarchical categorization. *Journal of the Experimental Analysis of Behavior*, *78*, 433-447.

Gross, A. C., & Fox, E. J. (2009). Relational frame theory: An overview of the controversy. *The Analysis of Verbal Behavior, 25,* 87-98.

Harlow, H. E. (1959). Learning set and error factor theory. In S. Koch (Ed.), *Psychology: A study of a science* (Vol. 2, pp. 492-537). New York, NY: McGraw-Hill.

Harmon, K., Strong, R., & Pasnak, R. (1982). Relational responses in tests of transposition with rhesus monkeys. *Learning and Motivation, 13,* 495-504.

Hayes, S. C. (1987). Upward and downward continuity: It's time to change our strategic assumptions. *Behavior Analysis, 22,* 3-6.

Hayes, S. C. (1989a). Nonhumans have not yet shown stimulus equivalence. *Journal of the Experimental Analysis of Behavior, 51,* 385-392.Hayes, S. C. (Ed.). (1989b). *Rule-governed behavior: Cognition, contingencies, and instructional control*. New York,NY: Plenum.

Hayes, S. C. (1992). Verbal relations, time, and suicide. In S. C. Hayes & L. J. Hayes (Eds.), *Understanding verbal relations* (pp. 109-118). Reno, NV: Context Press.

Hayes, S. C. (2002). Acceptance, mindfulness, and science. *Clinical Psychology: Science and Practice, 9,* 101-106.

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

Hayes, S. C., Brownstein, A. J., Haas, J. R., & Greenway, D. E. (1986). Instructions, multiple schedules, and extinction: distinguishing rule-governed from scheduled-controlled behavior. *Journal of the Experimental Analysis of Behavior, 46,* 137-147.

Hayes, S. C., Brownstein, A. J., Zettle, R. D., Rosenfarb, I., & Korn, Z. (1986). Rule-governed behavior and sensitivity to changing consequences of responding. *Journal of the Experimental Analysis of Behavior, 45,* 237-256.

Hayes, S. C., Fox, E., Gifford, E. V., Wilson, K. G., Barnes-Holmes, D., & Healy, O. (2001). Derived relational responding as learned behavior. In S. C. Hayes, D. Barnes-Holmes, & B. Roche (Eds.), *Relational frame theory: A post Skinnerian account of human language and cognition* (pp. 21–49). New York, NY: Plenum.

Hayes, S.C., & Long, D. (2013). Contextual behavioral science, evolution, and scientific epistemology. In S. Dymond & B. Roche (Eds.), *Advances in Relational Frame Theory: Research & Application*. Oakland, CA: New Harbinger.

Hayes, S. C., Strosahl, K., & Wilson, K. G. (1999). *Acceptance and commitment therapy: An experiential approach to behavior change*. New York, NY: Guilford.

Heagle, A. I., & Rehfeldt, R. A. (2006). Teaching perspective-taking skills to typically developing children through derived relational responding. *Journal of Early and Intensive Behavior Intervention, 3,*1-34.

Healy, O., Barnes, D., & Smeets, P. M. (1998). Derived relational responding as an operant: The effects of between-session feedback. *The Psychological Record, 48,* 511-536.

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.

Horne, P. J., & Erjavec, M. (2007). Do infants show generalized imitation of gestures? *Journal of Experimental Analysis of Behavior, 87,* 63-87.

Hughes, S., & Barnes-Holmes, D. (2014). Associative concept learning, stimulus equivalence, and relational frame theory: Working out the similarities and differences between human and non-human behavior. *Journal of Experimental Behavior Analysis, 101,* 156-160.

Hughes, S., Barnes-Holmes, D., & Vahey, N. (2012). Holding on to our functional roots when exploring new intellectual islands: A voyage through implicit cognition research. *Journal of Contexual Behavioral Science, 1,* 17-38.

Hughes. S., De Houwer, J., & Barnes-Holmes, D. (2014). On how contextual behavioral science may contribute to the study of evaluative conditioning. *Unpublished manuscript*

Hull, C. L. (1934). The concept of the habit-family hierarchy and maze learning: Part 1. *Psychological Review, 41,* 33-52.

Hyland, J.M., Smyth, S., O'Hora, D.P. & Leslie, J.C. (2014). The effect of before and after instructions on the speed of sequential responding. *The Psychological Record, 64,* 311-319.

Jenkins, J. J., & Palermo, D. S. (1964). Mediation processes and the acquisition of linguistic structure. In U. Bellugi & R. Brown (Eds.), *The acquisition of language*. Monographs of the Society for Research in Child Development, 29, (pp.141-169).

Kastak, C. R., & Schusterman, R. J. (2002). Sea lions and equivalence: Expanding classes by exclusion. *Journal of the Experimental Analysis of Behavior, 78,* 449-465.

Leigland, S. (1997). Is a new definition of verbal behavior necessary in the light of derived relational responding? *The Behavior Analyst, 20,* 3-10.

Leonhard, C., & Hayes, S. C. (1991, May). *Prior inconsistent testing affects equivalence responding*. Paper presented at the meeting of the Association for Behavior Analysis, Atlanta.

Lieving, G. A., & Lattal, K. A. (2003). Recency, repeatability, and reinforcer retrenchment: An experimental analysis of resurgence. *Journal of the Experimental Analysis of Behavior, 80,* 217-233.

Lionello-DeNolf, K. M. (2009). The search for symmetry: 25 years in review. *Learning & Behavior, 37,* 188-203.

Lipkens, G., & Hayes, S. C. (2009). Producing and recognizing analogical relations. *Journal of the Experimental Analysis of Behavior, 91,* 105–126.

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.

Lowe, C. F. (1979). Determinants of human operant behaviour. In M. Zeiler & P. Harzem (Eds.), *Advances in the analysis of behavior: Volume 1. Reinforcement and the organisation of behavior* (pp.159-192). Chichester, United Kingdom: Wiley.

Luciano, C., Gómez-Becerra, I., & Rodríguez-Valverde, M. (2007). The Role of multiple -exemplar training and naming in establishing derived equivalence in an infant. *Journal of Experimental Analysis of Behavior, 87,* 349-365.

Luciano, C., Rodrigquez, M., Manas, I., & Ruiz, F. (2009). Acquiring the earliest relational operants: Coordination, difference, opposition, comparison, and hierarchy. In R. A. Rehfeldt & Y. Barnes-Holmes (Eds.), *Derived relational responding: Applications for learners with autism and other developmental disabilities* (pp. 149–172). Oakland, CA: New Harbinger.

Luciano, C., Valdivia-Salas, S., Ruiz, F. J., Rodríguez-Valverde, M., Barnes-Holmes, D., Dougher, M. J., & Gutierrez, O. (2013). Extinction of aversive eliciting functions as an analog of exposure to conditioned fear: Does it alter avoidance responding? *Journal of Contextual Behavioral Science, 2,* 120–134.

Luciano, C., Valdivia-Salas, S., Ruiz, F. J., Rodríguez-Valverde, M., Barnes-Holmes, D., Dougher, M. J.,…Gutierrez-Martínez, O. (2014). Effects of an acceptance/defusion intervention on experimentally induced generalized avoidance: A laboratory demonstration. *Journal of the Experimental Analysis of Behavior, 101,* 94-111.

McGinty, J., Ninness, C., McCuller, G., Rumph, R., Goodwin, R., Kelso, G.,…Kelly, E. (2012). Training and deriving precalculus relations: A small group web-interactive approach. *The Psychological Record, 62,* 225–242**.**

McHugh, L., Barnes-Holmes, Y., & Barnes-Holmes, D. (2004). Perspective-taking as relational responding: A developmental profile. *The Psychological Record, 54,* 115-144.

McHugh, L., Barnes-Holmes, Y., Barnes-Holmes, D., & Stewart, I. (2006). Understanding false belief as generalized operant behavior. *The Psychological Record, 56,* 341-364.

McHugh, L., Barnes-Holmes, Y., & Barnes-Holmes, D. (2007). Deictic relational complexity and the development of deception. *The Psychological Record, 57,* 517-531.

McHugh, L., Barnes-Holmes, Y., Barnes-Holmes, D., Stewart, I., & Dymond, S. (2007). Deictic relational complexity and the development of deception. *The Psychological Record*, *57,*517-531.

McHugh, L., & Stewart, I. (2012). *The self and perspective taking: Contributions and applications from modern behavioral science*. Oakland, CA: New Harbinger .

McIlvane, W. J. (2014). Associative concept learning in animals by Zentall, Wasserman, and Urcuioli: A commentary. *Journal of the Experimental Analysis of Behavior, 101,* 161 -164.

McIlvane, W. J., Serna, R. W., Dube, W. V., & Stromer, R. (2000). Stimulus control topography coherence and stimulus equivalence: Reconciling test outcomes with theory. In J. C. Leslie & D. E. Blackman (Eds.), *Experimental and applied analysis of human behavior* (pp. 85-110). Reno, NV: Context Press.

Meltzoff, A. N. (2005). Imitation and other minds: The "Like Me" hypothesis. In S. Hurley & N.Chater (Eds.), *Perspectives on imitation: From cognitive neuroscience to social science* (Vol. 2,pp. 55-77). Cambridge, MA: MIT Press.

Mischel, W., & Ayduk, O. (2004). Willpower in a cognitive-affective processing system: The dynamics of delay of gratification. In R. F., Baumeister & K. D. Vohs. (Eds). *Handbook of self-regulation: Research, theory, and applications* (p. 99-129). New York, NY: Guilford.

Munnelly, A. Freegard, G. & Dymond, S. (2013). Constructing relational sentences: Establishing arbitrarily applicable comparative relations with the Relational Completion Procedure. *The Psychological Record, 63,* 751-768.

Munnelly, A., Martin, G., Dack, C., Zedginidze, A., & McHugh, L. (in press). The transfer of social exclusion and inclusion functions through derived stimulus relations. *Learning & Behavior*.

Murphy, C., & Barnes-Holmes, D. (2009). Derived more–less relational mands in children diagnosed with autism. *Journal of Applied Behavior Analysis, 42,* 253-268.

Murphy, C., Barnes-Holmes, D., & Barnes-Holmes, Y. (2005). Derived manding in children with autism: Synthesizing Skinner's verbal behavior with relational frame theory. *Journal of Applied Behavior Analysis, 38),* 445-462.

Murphy, G. L. (2002). *The big book of concepts*. Cambridge, MA: MIT Press.

Neuringer, A. (2002). Operant variability: Evidence, functions, and theory. *Psychonomic Bulletin and Review, 9,* 672–705.

Ninness, C., Barnes-Holmes, D. Rumph, R., McCuller, G., Ford, A., Payne, R., . . .Elliott, M. (2006). Transformation of mathematical and stimulus functions. *Journal of Applied Behavior Analysis,* 39, 299-321.

Ninness, C., Dixon, Barnes-Holmes, D., Rehfeldt, R.A., Rumph, R., McCuller, G., . . .Ninness, S. (2009). Deriving and constructing reciprocal trigonometric relations: A web -interactive training approach. *Journal of Applied Behavior Analysis, 43,* 191-208.

O'Connor, J., Barnes-Holmes, Y., & Barnes-Holmes, D. (2011). Establishing contextual control over symmetry and asymmetry performances in typically-developing children and children with autism. *The Psychological Record, 61,* 287-312.

O’Connor, J., Rafferty, A., Barnes-Holmes, D, & Barnes-Holmes, Y. (2009). The role of verbal behavior, stimulus nameability, and familiarity on the equivalence performances of autistic and normally developing children. *The Psychological Record, 59,* 53-74.

O’Hora, D., & Barnes-Holmes, D. (2004). Instructional control: Developing a relational frame analysis. *International Journal of Psychology and Psychological Therapy, 4,*263-284.

O’Hora, D., Barnes-Holmes, D., Roche, B., & Smeets, P. M. (2004). Derived relational networks and control by novel instructions: A possible model of generative verbal responding. *The Psychological Record, 54,* 437-460.

O’Hora, D., Pelaez, M., Barnes-Holmes, D., & Amesty, L. (2005). Derived relational responding and human language: Evidence from the WAIS-III. *The Psychological Record, 55,* 155-174.

O'Hora, D., Pelaez, M., Barnes-Holmes, D., Rae, G., Robinson, K., & Chaudhary, T. (2008). Temporal relations and intelligence: Correlating relational performance with performance on the WAIS-III. *The Psychological Record, 58,* 569-584.

O’Hora, D., Roche, B., Barnes-Holmes, D., & Smeets, P. M. (2002). Response latencies to multiple derived stimulus relations: Testing two predictions of relational frame theory. *The Psychological Record, 52,* 51–76.

O’Toole, C., & Barnes-Holmes, D. (2009). Three chronometric indices of relational responding as predictors of performance on a brief intelligence test: The importance of relational flexibility. *The Psychological Record, 59,* 119-132.

Pankey, J., & Hayes, S. C. (2003). Acceptance and commitment therapy for psychosis. *International Journal of Psychology and Psychological Therapy, 3,* 311-328.

Pelaez, M., Gewirtz, J. L., Sanchez, A., & Mahabir, N. M. (2000). Exploring stimulus equivalence formation in infants. *Behavior Development Bulletin, 9,* 20-25.

Pérez, V., García, A., & Gómez, J. (2011). Facilitation of the equivalence-equivalence response. *Psicothema, 23,* 407-414.

Persicke, A., Tarbox, J., Ranick, J., & St. Clair, M. (2012). Establishing metaphorical reasoning in children with autism. *Research in Autism Spectrum Disorders, 6),* 913–920.

Pilgrim, C., Click, R., & Galizio, M. (2011). A developmental analysis of children’s equivalence-class formation and disruption. *Acta de Investigacion Psicologica, 1,* 55-76.

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

Posner, M. I., & Rothbart, M. K. (2007). Research on attention networks as a model for the integration of psychological science. *Annual Review of Psychology, 58,* 1-23.

Quinones, J. L., & Hayes, S. C. (2014). Relational coherence in ambiguous and unambiguous relational networks. *Journal of the Experimental Analysis of Behavior*, *101,* 76–93.

Reese, H. W. (1968). *The perception of stimulus relations: Discrimination learning and transposition*. New York, NY: Academic Press.

Rehfeldt, R. A., & Barnes-Holmes, Y. (2009). *Derived relational responding: Applications for learners with autism and other developmental disabilities*. Oakland, CA: New Harbinger.

Rehfeldt, R. A., Dillen, J. E., Ziomek, M. M., & Kowalchuk, R. E. (2007). Assessing relational learning deficits in perspective-taking in children with high functioning autism spectrum disorder. *The Psychological Record, 57,* 23–47.

Reilly, T., Whelan, R., & Barnes-Holmes, D. (2005). The effect of training structure on the latency responses to a five-term linear chain. *The Psychological Record, 55,* 233–249.

Roche, B., & Barnes, D. (1996). Arbitrarily applicable relational responding and human sexual categorization: A critical test of the derived difference relation. *The Psychological Record, 46,* 451-475.

Roche, B., & Barnes, D., & Smeets, P. M. (1997). Incongruous stimulus pairing contingencies and conditional discrimination training: Effects on relational responding. *Journal of the Experimental Analysis of Behavior*, *68,* 143-160.

Roche, B., Barnes-Holmes, Y., Barnes-Holmes, D., Stewart, I., & O'Hora, D. (2002). Relational frame theory: A new paradigm tor the analysis of social behavior. *The Behavior Analyst, 25,* 75-91.

Roche, B. T., Kanter, J. W., Brown, K. R., Dymond, S., & Fogarty, C. C. (2008). A comparison of "direct" versus "derived" extinction of avoidance responding. *The Psychological Record, 58,* 443-464.

Roche, B., Linehan, C., Ward, T., Dymond, S., & Rehfeldt, R. (2004). The unfolding of the relational operant: A real-time analysis using electroencephalography and reaction time measures. *International Journal of Psychology & Psychological Therapy, 4,* 587-603.

Rodríguez-Valverde, M., Luciano, C., & Barnes-Holmes, D. (2009). Transfer of aversive respondent elicitation in accordance with equivalence relations. *Journal of the Experimental Analysis of Behavior, 92,* 85-111.

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.

Sautter, R. A., & LeBlanc, L. A. (2006). Empirical applications of Skinner's analysis of verbal behavior with humans. *The Analysis of Verbal Behavior, 22,* 35-48.

Schlinger, H. D. (1993). Separating discriminative and function-altering effects of verbal stimuli. *The Behavior Analyst, 16,* 9-23.

Schlinger, H. D. (2008). The long goodbye: Why B. F. Skinner's *Verbal Behavior* is alive and well on the 50th anniversary of its publication. *The Psychological Record, 58,* 329-337.

Schlinger, H. D. (2010). The impact of Skinner's *Verbal Behavior*: A response to Dymond and Alonso-Alvarez. *The Psychological Record, 60,* 361-368.

Schlinger, H., & Blakely, E. (1987). Function-altering effects of contingency-specifying stimuli. *The Behavior Analyst, 10,* 41–45.

Shimoff, E., Catania, A. C., & Matthews, B. A. (1981). Uninstructed human responding: Sensitivity of low-rate performance to schedule contingencies. *Journal of the Experimental Analysis of Behavior*, *36*, 207–220.

Skinner, B. F. (1957). *Verbal behavior*. Englewood Cliffs, NJ: Prentice Hall.

Skinner, B. F. (1969). *Contingencies of reinforcement: A theoretical analysis*. Englewood Cliffs, NJ: Prentice-Hall.

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, MA: Authors Cooperative.

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

Sidman, M. (2009). Equivalence relations and behavior: An introductory tutorial. *The Analysis of Verbal Behavior, 25,* 5–17.

Slattery, B., & Stewart, I. (2014). Hierarchical classification as relational framing. *Journal of the Experimental Analysis of Behavior, 101*, 61-75.

Slattery, B., Stewart, I., & O'Hora, D. (2011). Testing for transitive class containment as a feature of hierarchical classification. *Journal of the Experimental Analysis of Behavior, 96*, 243-260.

Smeets, P. M., & Barnes-Holmes, D. (2005). Establishing equivalence classes in preschool children with many-to-one and one-to-many training protocols. *Behavioural Processes, 69,* 281-293.

Smeets, P. M., Barnes-Holmes, Y., Akpinar, D., & Barnes-Holmes, D. (2003). Reversal of equivalence relations. *The Psychological Record, 53,* 91-120.

Smyth, S., Barnes-Holmes, D., & Forsyth, J. P. (2006). A derived transfer of simple discrimination and self-reported arousal functions in spider fearful and non-spider fearful participants. *Journal of the Experimental Analysis of Behavior, 85,* 223-246.

Steele, D. L., & 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. (2004). Relational Frame Theory and analogical reasoning: Empirical investigations. *International Journal of Psychology and Psychological Therapy, 4,* 241-262.

Stewart, I., Barrett, K., McHugh, L., Barnes-Holmes, D., & O'Hora, D. (2013). Multiple contextual control over non-arbitrary relational responding and a preliminary model of pragmatic verbal analysis. *Journal of the Experimental Analysis of Behavior, 100,* 174 -186.

Stewart, I., & Roche, B. (2013). Relational frame theory: An overview. In S. Dymond & B. Roche (Eds.), *Advances in Relational Frame Theory: Research and application.* Oakland, CA: New Harbinger.

Tomanari, G. Y., Sidman, M., Rubio, A. R., & Dube, W. V. (2006). Equivalence classes with requirements for short response latencies. *Journal of the Experimental Analysis of Behavior, 85,* 349–369.

Urcuioli, P. J. (2008). Associative symmetry, "anti-symmetry", and a theory of pigeons' equivalence-class formation. *Journal of the Experimental Analysis of Behavior, 90,* 257–282.

Vaughan, M. E. (1985). Repeated acquisition in the analysis of rule-governed behavior. *Journal of the Experimental Analysis of Behavior, 44,* 175-184.

Vervoort, E., Vervliet, B., Bennett, M., & Baeyens, F. (2014). Generalization of human fear acquisition and extinction within a novel arbitrary stimulus category. PLoS ONE 9(5):

e96569. doi:10.1371/journal.pone.0096569

Vilardaga, R. (2009). A relational frame theory account of empathy. *The International Journal of Behavioral Consultation and Therapy, 5,* 178-184.

Vilardaga, R., Estévez, A., Levin, M. E., & Hayes, S. C. (2012). Deictic relational responding, empathy, and experiential avoidance as predictors of social anhedonia: Further contributions from relational frame theory. *The Psychological Record, 62,* 409–432.

Vilardaga, R. & Hayes, S. C. (2012). A contextual behavioral approach to pathological altruism. In B. Oakley, A. Knafo, G. Madhavan, & D. S. Wilson (Eds). *Pathological altruism* (pp. 31-48). New York, NY: Oxford University Press.

Villatte, M., Monestès, J. L., McHugh, L., Freixa i Baqué, E., & Loas, G. (2008). Assessing deictic relational responding in social anhedonia: A functional approach to the development of Theory of Mind impairments. *International Journal of Behavioral Consultation and Therapy, 4,* 360-373.

Villatte, M., Monestès, J. L., McHugh, L., Freixa i Baqué, E., & Loas, G. (2010). Adopting the perspective of another in belief attribution: Contribution of Relational Frame Theory to the understanding of impairments in schizophrenia. *Journal of Behavior Therapy and Experimental Psychiatry, 41,* 125-134.

Vitale, A., Barnes-Holmes, Y., Barnes-Holmes, D., & Campbell, C. (2008). Facilitating responding in accordance with the relational frame of comparison: Systematic empirical analyses. *The Psychological Record, 58,* 365-390.

Vitale, A., Campbell, C., Barnes-Holmes, Y., & Barnes-Holmes, D. (2012). Facilitating responding in accordance with the relational frame of comparison II: Methodological analysis. *The Psychological Record*, *62,* 663–675.

Walsh, S., Horgan, J., May, R. J., Dymond, S., & Whelan, R. (2014). Facilitating relational framing in children and individuals with developmental delay using the Relational Completion Procedure. *Journal of the Experimental Analysis of Behavior, 101,* 51-60.

Wang, T., McHugh, L., & Whelan, R. (2012). A test of the discrimination account in equivalence class formation. *Learning and Motivation, 43,* 8–13.

Watt, A., Keenan, M., Barnes, D., & Cairns, E. (1991). Social categorization and stimulus equivalence. *The Psychological Record, 41,* 33-50.

Weil, T. M., Hayes, S. C., & Capurro, P. (2011). Establishing a deictic relational repertoire in young children. *The Psychological Record, 61,* 371-390.

Whelan, R., & Barnes-Holmes, D. (2004a). Empirical models of formative augmenting in accordance with the relations of same, opposite, more-than, and less-than. *International Journal of Psychology and Psychological Therapy, 4,* 285-302.

Whelan, R., & Barnes-Holmes, D. (2004b). The transformation of consequential functions in accordance with the relational frames of same and opposite. *Journal of the Experimental Analysis of Behavior, 82,* 177-195.

Whelan, R., Barnes-Holmes, D., & Dymond, S. (2006). The transformation of consequential functions in accordance with the relational frames of more-than and less-than. *Journal of the Experimental Analysis of Behavior, 86,* 317-335.

Whelan, R., Cullinan, V., & O'Donovan, A. (2005). Derived same and opposite relations produce association and mediated priming. = Mismas Relaciones derivadas y Opuestas Producen la Asociación y la Preparación Mediada. *International Journal of Psychology & Psychological Therapy, 5,* 247-264.

Whelan, R., & Schlund, M. W. (2013). Reframing relational frame theory research: Gaining a new perspective through the application of novel behavioral and neurophysiological methods. In S. Dymond & B. Roche (Eds.), *Advances in RelationalFrame Theory: Research and application*. Oakland, CA: New Harbinger.

Wilson, K. G., & Hayes, S. C. (1996). Resurgence of derived stimulus relations. *Journal of the Experimental Analysis of Behavior, 66*, 267-281.

Wray, A.M., Dougher, M.J., Hamilton, D.A., & Guinther, P.M. (2012). Examining the reinforcing properties of making sense: A preliminary Investigation. *The Psychological Record, 62,* 599-622.

Wulfert, E., & Hayes, S. C. (1988). Transfer of a conditional ordering response through conditional equivalence classes. *Journal of the Experimental Analysis of Behavior, 50,* 125–144.

Yamazaki, Y., Saiki, M., Inada, M., Iriki, A., & Watanabe, S. (*in press*). Transposition and its generalization in common marmosets. *Journal of Experimental Psychology: Animal Learning and Cognition*.

Yanchar, S. C., & Slife, B. D. (1997). Pursuing unity in a fragmented psychology: Problems and prospects. *Review of General Psychology, 1,* 235–255.

Zentall, T. R., Wasserman, E. A., & Urcuioli, P. J. (2014). Associative concept learning in animals. *Journal of the Experimental Analysis of Behavior, 101,* 130-151.