

RESEARCH ARTICLE

Verbal behavior development theory and relational frame theory: Reflecting on similarities and differences

Maithri Sivaraman¹ | Dermot Barnes-Holmes² | R. Douglas Greer³ |
Daniel M. Fienup³ | Herbert Roeyers⁴

¹School of Psychology, Cardiff University, Cardiff, Wales, United Kingdom

²School of Psychology, Ulster University, Coleraine, N. Ireland, United Kingdom

³Department of Health and Behavior Studies, Teachers College, Columbia University, New York, USA

⁴Department of Experimental, Clinical, and Health Psychology, Ghent University, Ghent, Belgium

Correspondence

Maithri Sivaraman, School of Psychology, Cardiff University, UK.
Email: sivaramanm@cardiff.ac.uk

Editor-in Chief: Mark Galizio

Handing Editor: Mark Galizio

Abstract

Relational frame theory and verbal behavior development theory are two behavior-analytic perspectives on human language and cognition. Despite sharing reliance on Skinner's analysis of verbal behavior, relational frame theory and verbal behavior development theory have largely been developed independently, with initial applications in clinical psychology and education/development, respectively. The overarching goal of the current paper is to provide an overview of both theories and explore points of contact that have been highlighted by conceptual developments in both fields. Verbal behavior development theory research has identified how behavioral developmental cusps make it possible for children to learn language incidentally. Recent developments in relational frame theory have outlined the dynamic variables involved across the levels and dimensions of arbitrarily applicable relational responding, and we argue for the concept of mutually entailed orienting as an act of human cooperation that drives arbitrarily applicable relational responding. Together these theories address early language development and children's incidental learning of names. We present broad similarities between the two approaches in the types of functional analyses they generate and discuss areas for future research.

KEYWORDS

incidental bidirectional naming, cooperation, naming, orienting, relational frame theory, verbal behavior development

Fifty years ago, Murray Sidman published the seminal article on what subsequently became known as stimulus equivalence (Sidman, 1971). When a series of related conditional discriminations are trained, equivalence relations emerge when the stimuli involved in those discriminations often become related to each other in ways that were not explicitly trained (e.g., when A1-B1 and B1-C1 are trained, C1-A1 emerges without explicit training). It was over 10 years before it became clear that the stimulus equivalence phenomenon may be shown with relative ease in humans but far less readily, if at all, in nonhuman animals (Sidman et al., 1982), including higher primates. Subsequent research published 20 years later still indicated that even language-trained chimpanzees failed to show the most rudimentary elements of equivalence class formation (Dugdale & Lowe, 2000).

The connection between stimulus equivalence and *symbolic* relations in human language thus became the focus of both the empirical research and the sometimes-heated conceptual debate at that time (see Sidman, 1994, and specifically the written exchange between Sidman and Day). In broad terms, Sidman and colleagues argued that equivalence relations provided the basis for symbolic relations (see Sidman, 1994, for a book length review; see also Sidman, 2000), whereas other researchers argued that equivalence was explained by symbolic relations, typically involved in learning to name objects and events (e.g., Horne & Lowe, 1996). A third perspective, provided by relational frame theory (RFT; Hayes & Hayes, 1989), suggested that equivalence relations and symbolic relations in natural language were essentially synonymous, functionally

This is an open access article under the terms of the [Creative Commons Attribution](https://creativecommons.org/licenses/by/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2023 The Authors. *Journal of the Experimental Analysis of Behavior* published by Wiley Periodicals LLC on behalf of Society for the Experimental Analysis of Behavior.

speaking. And the explanation for both was to be found in a history of generalized operant learning defined as arbitrarily applicable relational responding (AARR; e.g., Hayes et al., 2001).

Researchers studying stimulus equivalence, naming, and RFT have tended to work largely independently with distinct research agendas. Verbal behavior development theory (VBDT) focuses on the study of how naming from exposure alone is acquired in children and on the different phases leading to that development (first the listener responses, then the speaker responses; Greer & Ross, 2008; see also Pérez-Gonzalez et al., 2014). Other naming theorists have investigated how bidirectional naming facilitates equivalence, categorization, and problem solving (see Miguel, 2018, for a review). RFT researchers have studied the concepts of relational frames or different patterns of derived relations to build a general functional-analytic approach to human language and cognition (e.g., see Barnes-Holmes et al., 2018, for a recent review). Other researchers have focused on what has become known as equivalence-based instruction (e.g., Carr et al., 2000; Fienup & Brodsky, 2020).

The primary purpose of the current article is to begin to find points of mutual interest and contact across the different research approaches to human language and how we might begin to work together in a genuinely collaborative way, hopefully to the benefit of all (see also Fienup, 2018, for an editorial on the future of verbal behavior research). In the next sections, we will focus on the relationship between recent work in VBDT and recent developments in RFT. In doing so, we are not excluding the importance of other approaches, such as equivalence or equivalence-based instruction, but trying to stay focused on two areas where there seems to be a high degree of overlap. For example, bidirectional naming was always seen to be critically important or central to derived relations for both naming theorists¹ and RFT researchers. Although early treatments tended to emphasize different views (e.g., whether equivalence explained symbolic relations or symbolic relations explained equivalence), more recent developments suggest that such theoretical differences may be less important going forward.

More recently, the emphasis on the distinction between unidirectional and bidirectional naming (and incidental naming) and a focus on the evolution of language (Pohl et al., 2018) seems to be in line with a recent focus on linking RFT more closely to the evolution of cooperation in humans (Hayes & Sanford, 2014). Furthermore, conceptual developments in RFT, which have provided a general framework (hyperdimensional multilevel framework;

Barnes-Holmes et al., 2020) and a dynamic unit of analysis (relating, orienting, evoking, and motivation [ROE-M]; Barnes-Holmes & Harte, 2022) have served to highlight clear points of contact and overlap between the analysis of different forms of bidirectional naming (BiN) and different levels and dimensions of derived relating within RFT. Each of these concepts will be discussed in detail in upcoming sections of the paper. Before continuing, we should emphasize that in drawing on recent conceptual developments in RFT in the context of the current article, we are not arguing that RFT only *now* connects more readily with naming research. The potential links between RFT and naming were always present in the literature, going back to the earliest RFT studies (e.g., Barnes et al., 1990; Lipkens et al., 1993). However, recent conceptual developments in both RFT and VBDT have served to bring the potential overlap into relatively sharp focus, at least for the current authors, and the present article is just one attempt to explore and hopefully exploit this potential overlap.

A VERBAL BEHAVIOR DEVELOPMENTAL THEORY OF BIDIRECTIONAL NAMING

Verbal development research began out of the need to build a more complete *science of teaching* based on the science of behavior (Barrett et al., 1991; Greer, 2002; Skinner, 1954). Skinner's (1957) description of verbal (speaker) operants, multiple stimulus control, and antecedent and postcedent verbal stimuli has proved useful in following or constructing algorithms for solving complex problems (e.g., Williams & Greer, 1993). Incorporating Skinner's (1957) work necessitated a program of research in how verbal behavior *develops* as a function of a sequence of children's experiences and how that development affects pedagogy (Greer, 1991, 2002) and curriculum construction (Fienup & Brodsky, 2020; Williams & Greer, 1993). A behavior science of language seeks to identify how language repertoires are acquired by children including, for example, learning to respond as a listener and speaker from observing others saying the names of objects and events in the world. This phenomenon and the sequence of critical steps are called verbal developmental cusps and will be discussed in detail below.

Verbal behavior development: From data to theory

Skinner (1957) provided a framework for studying verbal development when he suggested that listener and speaker repertoires are initially independent and later combine into a behavioral unit. Specifically, Skinner (1957) argued that "once a speaker also becomes a listener, the stage is

¹Although the current article focusses on RFT and VBDT, the focus on naming, particularly incidental naming in the latter, allows us to draw on naming research more generally, and this will be reflected at various points throughout the current article when we refer to naming studies from outside of VBDT. Indeed, we see this as a positive strategy toward achieving more collaboration among all of the researchers in behavior analysis who are focusing on human language development.

set for a drama in which one man plays several roles” (p. 472). VBDT researchers examined this phenomenon by identifying children (18 months to 12 years) who lacked specific repertoires (such as imitating the actions of others) that have been identified as verbal developmental cusps and tested protocols to establish them (Greer & Ross, 2008; Novak & Pelaez, 2004; Rosales-Ruiz & Baer, 1996, 1997). A cusp is defined as a behavior that allows a child to learn things that could not be learned before, learn faster, or learn in new ways. It also leads to new behaviors without direct training or programmed reinforcement. Examples of cusps include motor imitation, echoing, and incidental bidirectional naming. Ross and Greer (2003), for instance, trained five children to imitate novel motor actions and observed subsequent improvements in their vocal imitation (echoics) and vocal requests (mands). Researchers also demonstrated that some cusps, such as bidirectional naming, allow children to learn in new ways. Greer et al. (2011) showed that children who emitted listener and speaker responses simply by hearing the name of an object without programmed instruction could respond correctly during math instructional trials that only involved an instructor modeling the target behavior (e.g., the instructor modeled identifying the units, tens, and hundreds places and subsequently presented the student with a worksheet to identify the place value of the underlined digit). In contrast, children who did not demonstrate incidental BiN could not respond correctly on such trials without programmed reinforcement.

Categories of verbal developmental cusps

According to VBDT, a behavior is said to be a verbal developmental cusp if subsequent to its acquisition (a) a child can learn verbal repertoires they could not learn before the onset of the cusp, (b) learn these repertoires significantly faster (e.g., no prompts), or (c) learn in ways they could not before (e.g., from observation alone). These cusps fall into four categories: (a) preverbal foundational cusps, (b) listener response cusps, (c) speaker cusps, and (d) the cusps that show the joining of the listener and speaker between individuals and within the behavior of the individual. Preverbal foundation cusps include observing responses (orienting to or prolonged observing of faces and voices, two- and three-dimensional stimuli), generalized imitation, and echoing (Greer & Longano, 2010; see also Luciano & Polaino, 1986, which highlighted the importance of training orienting responses). Again, according to VBDT, preverbal foundational cusps set the stage for subsequent verbal cusps.

The listener cusps involve first instances of behavior coming under the stimulus control of spoken words (i.e., hear and then do; Choi et al., 2015). The initial listener cusps include responding to auditory stimuli including words and nonwords (e.g., phonemic discrimination),

responding to the words of a speaker as discriminative stimuli, and learning listener responses incidentally (pointing to an object after simply being exposed to an object-name relation, i.e., incidental unidirectional naming).

The tact operant is seen as a critical speaker cusp that is fundamental to the joining of the listener and speaker repertoires. VBDT data support Skinner’s (1957) account that tacts are acquired when behavior is reinforced by attention or praise (Eby & Greer, 2017; Schmelzkopf et al., 2017). In this, or other cases, a cusp is not just learning a single response; for example, teaching a child to say “tree” with arduous shaping and prompt procedures can result in the child saying “tree” on seeing a tree. However, if the child lacks the generalized tact operant, each tact will require arduous instruction. If a generalized tact operant is acquired, new incidences of tacts can be taught easily (see e.g., Eby & Greer, 2017; Schmelzkopf et al., 2017). When a generalized tact operant is acquired, only a small number of trials is needed to teach a novel tact. That is, tacts are maintained by naturally occurring reinforcement contingencies thereby giving the child opportunities to learn from the natural environment. VBDT focusses on training repertoires rather than individual responses. In addition, VBDT distinguishes between bidirectional operants (cusps) between and within individuals. Some examples of bidirectional cusps that involve the joining of the listener and speaker are discussed below. Before proceeding, however, it seems important to recognize that although the concept of verbal developmental cusps has proven useful in this area of research, the concept itself should remain somewhat provisional.

Bidirectional operants between individuals

The intersection of the speaker and listener is described as *speaker as own listener within the skin* by Skinner (1957, p. 476), reiterated as such by Horne and Lowe (1996, p. 240). However, in many cases, the role of the speaker and listener is shared between two individuals such as during any verbal episode. In such cases, when one individual speaks, the product of such behavior (i.e., verbal stimulus) affects both individuals as listeners, and it may also result in the second individual responding as a speaker, and so on (in VBDT this is referred to as a conversational unit, described as a verbal episode by Skinner, 1957). The speaker role is not limited to vocal communication and can also include nonspeech sounds, grimaces, frowns, or gestures. The acquisition of both listener and speaker behavior allows children to become more fully social (see Donley & Greer, 1993; Schmelzkopf et al., 2017).

It is important to note that the concept of a bidirectional operant could also be used to describe derived relational responding in a matching-to-sample procedure such as symmetry, and there is evidence that such performances may be observed in the absence of BiN (see for

example, Luciano et al., 2007). In this sense the concept of a bidirectional operant may be seen as broader than those described here which involve the joining of the listener and the speaker.

Bidirectional operants within individuals

Bidirectional operants acquired by young children may also occur in the absence of an “external” audience. These children engage in self-talk aloud, including say-then-do responding (see example below), during fantasy play (Lodhi & Greer, 1989). Incidences of self-talk can be directly observed when children rotate the roles of listener and speaker overtly as when a child acts as a speaker when interacting in solitary play with an anthropomorphic toy, as in saying, as a speaker, “Horsey, go to the barn,” and responding as a listener by moving the horse to the barn as identified by Lodhi and Greer (1989). Indeed, there may be some functional overlap between self-talk during fantasy play and say-and-do correspondence as identified in research by Paniagua and Baer (1982; see also Luciano et al., 2001).

Bidirectional naming and incidental language acquisition

At this point, it is important to note that there are several research programs, each focusing on different aspects of naming as defined by Horne and Lowe (1996), also called *common bidirectional naming* by Miguel (2016). Hawkins et al. (2018) proposed subclassifications of BiN based on (a) how the reinforcement of either speaker or listener responses results in the unreinforced emission of corresponding listener and speaker responses, respectively, and (b) how children learn names from experience without explicit training. VBBDT has focused on identifying how children may eventually come to acquire listener or speaker responses without the delivery of reinforcement by others. The VBBDT research agenda was largely driven by Hart and Risley’s (1995, 1999) classic study that found limited evidence of the parents’ conspicuous use of systematic reinforcement contingent upon speech, which could be seen as challenging for a behavior-analytic perspective of language development.

Many published studies on BiN (e.g., Horne et al., 2006; Lowe et al., 2002) focused on either the listener or speaker being taught (i.e., reinforced until mastery) and assessing for the untaught response under unreinforced conditions. Moreover, studies have investigated the relationship between naming and stimulus equivalence (Horne et al., 2004; Lowe et al., 2002; Lowe et al. 2005) including how precurrent or mediational verbal behavior may affect stimulus class formation (e.g., Jennings & Miguel, 2017; Lowenkron 1989, 1991; Ma et al., 2016; Miguel and Petursdottir, 2009). In

contrast, VBBDT researchers have focused on developing procedures testing the emission of *both listener and speaker responses* where *neither* has been directly reinforced (for a review, see Longano & Greer, 2014).

VBBDT research identified *multiple exemplar instruction* (MEI) as one intervention for training this form of incidental bidirectional naming (Inc-BiN; Fiorile & Greer, 2007; Gilic & Greer, 2011; Greer et al., 2005, Hawkins et al., 2009; Hotchkiss & Fienup, 2019; Luciano et al., 2007). The term *multiple exemplar training* (MET) is related to MEI (see Greer et al., 2017, for a description of the history of different types of MEI or MET; see also LaFrance and Tarbox, 2019, for a distinction in the current literature). The MEI intervention developed by VBBDT involved training children to respond across topographies of listener, speaker, and matching-to-sample trials across sets of stimuli until they demonstrated Inc-BiN (i.e., listener and speaker responses) to a novel set of stimuli without programmed reinforcement. Theoretically, the MEI involved explicitly reinforcing listener and speaker responses across exemplars until responses to novel stimuli emerged without reinforcement. Similarly, intensive tact instruction (Hotchkiss & Fienup, 2019; Schmelzkopf et al., 2017) and echoic training (Cao & Greer, 2019), which constitute elements of Inc-BiN, have also been shown to facilitate the establishment of this repertoire. In this sense, therefore, the effect of MEI is explained by appealing to histories of reinforcement across relevant exemplars, which establish new units of verbal behavior. Environmental contingencies may then act on these new units, thus producing novel behaviors (i.e., the behavioral process involved, therefore, is histories of reinforcement across exemplars).

Assessment of incidental bidirectional naming

The distinction between bidirectional naming as defined by Horne and Lowe (1996) and Inc-BiN is evident in the procedures used to assess the latter. In the assessment of Inc-BiN, researchers first identify pictures of stimuli that are novel and unfamiliar to a child. Once multiple sets of three to five stimuli are identified, the researchers present the child with each picture and simultaneously say the name of the picture; this is identified as the *naming experience*. Each of the five stimuli are presented four times for a block of 20 trials with no consequences for the child’s responses. Two hours later, the researchers present unreinforced probe trials to assess the accuracy during listener and speaker trials. First, the child is tested for the listener response where the researchers present arrays where the correct stimulus is presented along with two incorrect response stimuli with the instruction, “Point to the ___.” A child who demonstrates 80% accuracy in pointing to the stimuli is identified as demonstrating *incidental unidirectional naming* (Inc-UniN), meaning they acquire the names of things as a listener from exposure to

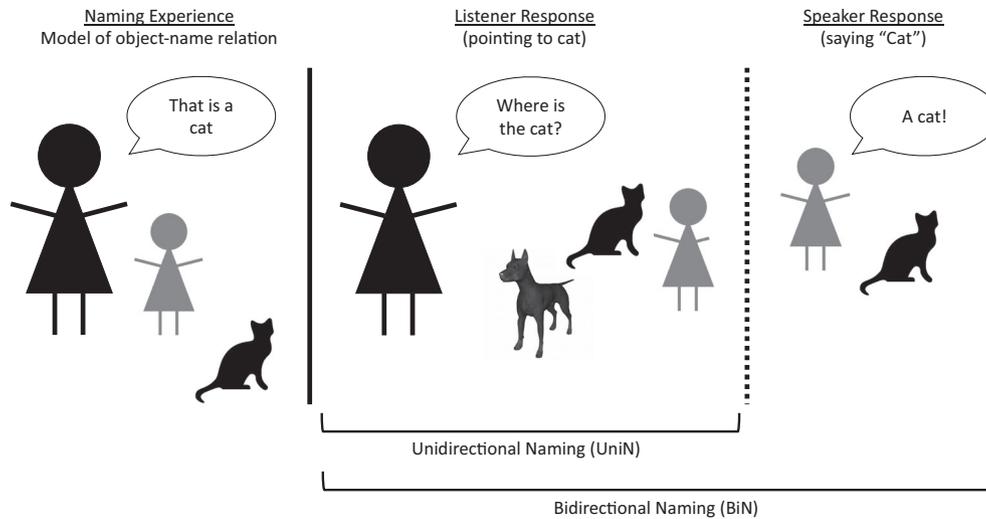


FIGURE 1 Incidental uni and bidirectional naming in a child

opportunities to observe caretakers say the names of things. Immediately after the listener probes are completed, and in the same session, researchers probe for the speaker responses. That is, each stimulus is presented two to four times, and the child is provided with the opportunity to say the name of the stimulus. If the child demonstrates 80% accuracy in speaking and listener responses, the child is identified as demonstrating Inc-BiN (see, for example, Morgan et al. 2021, for a detailed description of the assessment methods). Please see Figure 1 for an overview.

Incidental bidirectional naming is a continuum

Understanding how Inc-BiN is established has important implications for a science of behavior and of teaching. Children who do not acquire names incidentally learn very little from antecedent instruction (Greer et al., 2011), but once Inc-BiN is established they have the instructional history that results in learning in the absence of programmed reinforcement. Children who demonstrate Inc-UniN (i.e., listener naming but not speaker) need to be taught the speaker response but can emit the listener response without instruction (e.g., selecting, pointing to). Children who do not demonstrate Inc-UniN need to be taught both the listener and speaker responses. Children who demonstrate Inc-BiN can emit the untaught listener and speaker response from demonstration conditions alone (i.e., simply being exposed to object–name relations; Hrančuk et al., 2019). Furthermore, recent evidence shows that more complex verbal behavior emerges once Inc-BiN is established (Cahill & Greer, 2014; Greer & Du, 2015).

In this section, we described (a) the categories and types of cusps as defined by VBDT and how they contribute to verbal behavior development, (b) bidirectional operants (between and within individuals) that are seen as a critical cusp for engaging in verbal episodes with others,

and (c) a theoretical account of Inc-BiN and how children come to learn the names of things without instruction or programmed reinforcement. Although this is by no means a comprehensive account of all the empirical and conceptual analyses carried out by VBDT researchers, we have described the main aspects of the theory that lend to points of contact with RFT. These points of contact are discussed in the last section of the paper.

AN UPDATED RELATIONAL FRAME THEORY

In summarizing the RFT treatment of naming and subsequently comparing it with VBDT, it seems important to draw on a relatively up-to-date version of RFT rather than a version that was presented almost 20 years ago (Hayes et al., 2001). In this respect, two important advances in the theory appear to have been made. First, RFT now focuses on the evolution of cooperation in humans as a primary driver for AARR itself (Hayes & Sanford, 2014; Wilson et al., 2014). The focus on cooperation is relevant, but not critical, to the core argument we will make in the current article, so we will not dwell on it here. More critical is the second advance, which involves a proposed new framework that seeks to highlight or emphasize the dynamics involved across the levels and dimensions of AARR. Critically, this new framework focuses on the orienting and evoking functions of stimulus events, which as we shall see has important implications for an RFT treatment of naming.

A hyperdimensional multilevel framework

Researchers have recently proposed an overarching hyperdimensional multilevel (HDML) framework (Barnes-Holmes & Harte, 2022; Barnes-Holmes et al., 2017; Finn

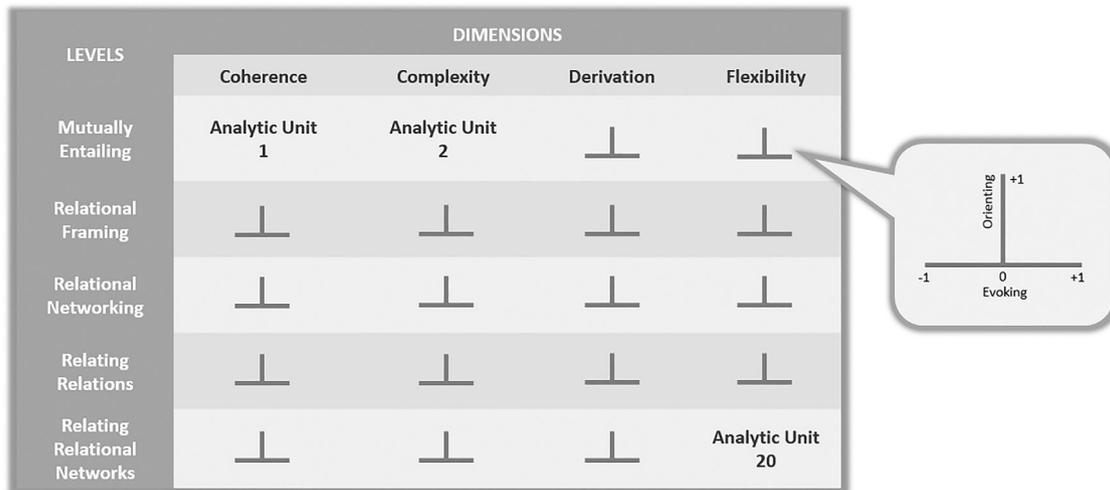


FIGURE 2 A Hyperdimensional multilevel framework consisting of 20 intersections between the dimensions and levels of arbitrarily applicable relational responding. The figure includes the properties of orienting and evoking.

et al., 2018; Harte et al., 2017, 2018) that summarizes how RFT approaches the *experimental analysis* of human language and cognition, with a view to revealing the complexity involved. The core idea of this development involves an attempt to capture a lot of what basic RFT researchers have been engaged in since the theory was first developed and to place that scientific behavior into a framework (i.e., the HDML) that emphasizes the dynamic nature of AARRing. One of the reasons for constructing such a framework is to help connect basic RFT research more directly with the concerns of applied researchers and practitioners.

The HDML framework specifies five levels and four dimensions of relational responding. The five levels are mutual entailing, relational framing (the simplest type of relational network), relational networking, relating relations, and relating relational networks. In addition, the HDML framework identifies the various dimensions of AARR that may be influenced by various contextual variables. Four such dimensions appear to be critically important (although others may emerge through future conceptual and empirical analyses) and are labeled as (a) coherence, (b) complexity, (c) derivation, and (d) flexibility. The framework currently consists of five levels and four dimensions of AARR that intersect to create 20 units (see Figure 2). The individual units may be useful in conceptualizing how to conduct experimental analyses of derived relational responding. In addition, the HDML framework explicitly incorporates three generic dimensions of transformations of function,²

²Transformation of stimulus functions was one of the three important properties of derived relational responding in the early RFT account. It was defined as a process in which the function of one stimulus in a derived relation alters the functions of another according to the relation between the two, without additional training (Hayes, 1991). Put simply, it was defined as the process by which stimuli come to acquire, change, or lose psychological properties. This account seemingly suggests that the transformation of stimulus functions occurs because of a derived relation between or among stimuli.

orienting, *evoking*, and *motivating* events. Orienting refers to the basic perceptual properties of a stimulus or event (including noticing or attending), and evoking refers to whether a perceived stimulus or event is appetitive, aversive, or relatively neutral. Motivating refers to the putative strength of motivational variables, which interact with orienting and/or evoking functions, and indeed relating, in a dynamic manner. In principle, any stimulus or stimulating event will possess these functions. Certainly, a stimulus cannot be defined as a stimulus if it does not produce at least some degree of orienting (i.e., a stimulus that is not perceived is not a stimulus).

In doing so, a general unit of conceptual analysis for RFT was proposed, which is referred to as the ROE-M (pronounced “roam”). The analytic unit conceptualizes individual derived relational responses as consisting of relating, orienting, evoking, and motivating events. In simple terms, relating refers to the myriad complex ways in which language-able humans can relate stimuli, as suggested by the 20 units of analysis within the HDML, and how the orienting and evoking functions of those stimuli both influence and are influenced by such relating. The reader is referred to Figure 2 for a visual representation of the HDML framework and Barnes-Holmes and Harte (2022) for more detail on the framework.

In recognizing the dynamic interplay among the properties of the ROE-M, it seems useful to conceptualize behavioral events for humans as involving a constant stream of relating (R), orienting (O), evoking (E), and motivating (M) events. For illustrative purposes, imagine a child is going into a garden looking for hidden Easter eggs and the caregiver says, “Watch out for a special golden egg that has gummy candies.” If the instruction is understood, it may be conceptualized as involving an instance of relating (e.g., relating the golden egg with candies), which may increase the likelihood that the child will orient toward any egg-like shape in the garden

followed by an appropriate evoked reaction, such as reaching forward and grabbing the egg (i.e., the functions of the egg have been transformed by the caregiver's instruction). In effect, the child's reaction to the egg is conceptualized as involving the elements of the ROE-M. As noted above, the ROE-M is conceptualized as a non-linear or dynamic unit of analysis and thus orienting and evoking may affect relating. Imagine, for example, that the child in the above example had not been given the instruction to look for a hidden Easter egg but instead had simply been asked to pick some flowers (i.e., the egg had been "hidden" by the caregiver as a "surprise"). When the child sees the egg, orienting and some evoking functions may occur, which then lead to some relating (e.g., the "oh, what's that—is it an Easter egg?" emitted publicly if the caregiver is present or perhaps privately if not).

The concept of mutually entailed orienting as the basis for shared intentionality

As noted earlier, in an updated RFT account, cooperation is theorized to be the primary initial driver of AARR. Hayes and Sanford (2014) described cooperation as consisting of "social referencing, joint attention, and perspective taking" (p. 122). In an effort to provide a more functional account of these terms, we argue here that one such act of cooperation involves *mutually entailed orienting*. This occurs when a child orients back and forth between a caregiver and an object or stimulus that the caregiver is orienting toward (this is not to be confused with orienting, per se, which may occur outside the context of a cooperative act). Specifically, infants have been shown to respond to social stimuli (e.g., adult voices, faces) from a very early age. For example, when a caregiver looks toward an object, say a Teddy bear, the infant looks at the caregiver, follows their gaze (or pointing) to the Teddy bear, and looks back at the caregiver (Tomasello, 1988). This sequence of coordinated attention between the infant, the caregiver, and an object, seems to occur constantly with multiple novel and nonnovel objects and events throughout the day.

In these episodes of very basic acts of cooperation, the infant's orienting may be seen as bidirectional in nature (i.e., back and forth between caregiver and object), and thus we might say that orienting toward the caregiver *entails* orienting toward the object, which entails orienting back toward the caregiver. In an updated version of RFT, we refer to this pattern as mutually entailed orienting (this behavior should not be confused with the RFT concept of mutual entailment, which would involve an instance of AARR, such as the infant symbolically relating an object and its name; see Barnes-Holmes & Harte, 2022). We make this claim because the behavior is deemed to be part and parcel of the evolution of a level and type of cooperation that is found in the human

species. Mutually entailed orienting should therefore be seen as a type of transgenerational phylogenetic behavior that is selected by reinforcement contingencies operating within the lifetime of the individual. In this sense, an updated version of RFT seeks increased scope in terms of linking directly with a modern evolutionary science (e.g., Wilson et al., 2014), which argues that evolution operates at multiple levels (e.g., genetic, cellular, symbolic, and cultural). Therefore, the historical basis for AARR does not begin with listening and speaking; it starts with one of the most basic of human cooperative acts (i.e., mutually entailed orienting). Furthermore, mutually entailed orienting provides the infant with an opportunity to continue interacting with the caregiver as a dyad, which likely serves as a reinforcer for continuing to engage in such acts of cooperation, thus supporting the historical context for the learning of AARR itself.³

The critical importance of mutually entailed orienting, in terms of survival value, cannot be underestimated because it allows caregivers to establish appetitive and aversive evoking functions for stimuli in the child's environment. From an evolutionary perspective, mutually entailed orienting increases the chances of survival and detecting danger. For example, if a caregiver shouts when the child approaches a dangerous stimulus (e.g., an insect with a powerful venom), that stimulus will likely acquire strong orienting and (aversive) evoking properties for the child. As a listening repertoire then develops, mutual entailing between specific sounds (i.e., words) and objects, such as dangerous insects, emerges. Gradually, therefore, a new response unit involving relating, orienting, and evoking is established for the child, and as noted previously we refer to this response unit as the ROE-M. This unit is seen as critical for the development of AARR, including, at least initially, naming as a listener repertoire.

In emphasizing the importance of cooperation as a driver of AARR and introducing the concept of mutually entailed orienting, the potential origins of contextual control over the transformation of functions becomes apparent. When an infant engages in mutually entailed orienting, even items that are simply oriented toward by the caregiver may become more valuable than other items in the environment and acquire relatively positive evoking (approach) functions for the infant. This likely serves as the basis for the shared intentionality observed in infants in a cooperative context such as an object-choice task where the researcher points to the object to be chosen (as opposed to apes who exhibit shared intentionality only in competitive contexts; Hare & Tomasello, 2004). That is, a basic transformation of function may occur as part and parcel of mutually entailed orienting (i.e., the evoking functions of an object may be

³The reader should note that engaging in mutually entailed orienting as part of a cooperative act with a caregiver does not necessarily preclude acts of competitiveness in infants with a noncaregiver (e.g., competing with a sibling for a toy).

transformed simply by orienting a child toward that stimulus). Thus, mutually entailed orienting is more accurately labeled mutually entailed orienting *and* evoking. The reader should note that the term *mutually entailed* (orienting and evoking) is employed to denote this type of human infant learning because it typically occurs in parallel with establishing a basic listener repertoire (e.g., a caregiver rarely engages a child in mutually entailed orienting and evoking without also emitting language-appropriate sounds, such as “Look, it’s a Teddy bear,” when orienting the child toward a Teddy bear). Employing the term mutual entailment (for orienting and evoking) also serves to highlight that this type of learning is viewed as emerging from an evolutionary history of cooperation found in the human species.

The present argument that mutually entailed orienting and evoking serve as the historical basis for AARR is broadly consistent with the findings of multiple longitudinal research studies (e.g., Carpenter et al., 1998; Mundy et al., 2007), which found that early word learning, both as a speaker and as a listener, is positively correlated with joint attention and orienting toward social stimuli. Furthermore, experimental investigations with infants (Slaughter & McConnell, 2003; Tenenbaum et al., 2014) revealed that following the gaze of the adult positively influenced infants’ learning of specific names for objects. These empirical findings are consistent with the arguments of Tomasello and Todd (1983), who suggested that infants learn the vast majority of words in their vocabulary through triadic social interactions. Similar findings, albeit smaller in number, have been reported in the behavior-analytic literature. Olaff and Holth (2020), for example, conditioned social stimuli as reinforcers and found improvement during probes for BiN in their participants; critically, the procedure for conditioning social stimuli as reinforcers required the participants to engage in orienting toward the researcher, shifting gaze toward the researcher, or making eye contact with the researcher. Thus, these behaviors occurred more frequently after the training phase and could have affected the positive BiN results they obtained. In addition, Maffei et al., (2014) reported increased emission of speaker behaviors (mands and tacts) following an intervention to establish conditioned reinforcement for faces in their participants; note that the participants were required to maintain eye contact with the face of the researcher as part of the training. All participants subsequently showed improvements in tacting during posttraining probes. Conversely, Harms (2020) implemented an intensive tact training protocol with six participants who exhibited little to no joint attention during pretraining probes. Following the intervention, five participants showed collateral improvements in receptive joint attention during posttraining probes. That is, these five participants were following the researcher’s gaze to the object that she was oriented toward and then looked back toward the researcher after the tact training intervention. It could thus

be argued that orienting toward faces was adventitiously reinforced during the tact training phase.

During the acts of cooperation involved in mutually entailed orienting/evoking, it is important to note that the caregiver does not necessarily become appetitive or aversive herself as a consequence of her reactions to the pleasurable and dangerous items in the environment. This may be the case at first—for example, if a child pulls away from or aggresses toward a caregiver when they shout at the child as a warning not to approach a dangerous object. However, an infant quickly learns to respond to the objects as being appetitive or aversive, and not the caregiver. Seven- to 15-month-old infants, for example, have been shown to be more likely to move away or toward a stranger after looking at the mother (Feinman, 1980), and procedures to establish such responding in young children have been tested (Pelaez et al., 2012; Sivaraman et al., 2022). In effect, mutually entailed orienting and evoking between the mother and numerous stimuli serves to establish the mother as a stimulus that transforms the functions of novel stimuli and events in the environment while maintaining generally appetitive functions for herself. In effect, a mother’s actions or behaviors may transform the functions of a novel stimulus, but the mother herself functions as a context for *limiting* the transformation of functions to that stimulus.

This control over (or limiting of) the transformation of functions could be seen as the basis for what is referred to as *Cfunc* control in RFT generally. This type of contextual control is seen as critical in selecting the specific functions that are transformed in any act of relating. For example, when an older child learns to relate the written word “candy” to actual candy they rarely attempt to eat the written word. Thus, the early cooperative acts involved in mutually entailed orienting and evoking in a sense provide the basis for the more sophisticated types of contextual control that are required as derived relational responding involving arbitrary stimuli is established in the child’s listening and speaking repertoires.

As listening and speaking are established through ongoing interactions between the child and its caregivers, extended cooperation further facilitates the adaptation of the species by allowing for more complex adaptations of the functional units, such as combinatorial entailment. For example, in young children, cooperation with the caregiver establishes the core functional unit of the ROE-M, which initially facilitates the emergence of listener behavior. That is, when the caregiver names a novel object, the child is subsequently able to orient toward the object (or point toward it/pick it up) when the name is uttered again, thereby exhibiting a very basic level of ROE-Ming (mutually entailed relating, orienting, evoking, and motivating). This level of ROE-Ming involves a relatively limited transformation of functions if orienting toward an object, pointing at it, or even picking it up are all functionally quite similar. For bidirectional naming to

emerge, the child not only orients toward the object (or points or picks it up) but also utters or vocalizes the sound that was heard during the mutually entailed orienting episode (i.e., when the object was named by the caregiver). This involves a more complex derived transformation of functions than mutually entailed orienting because the object is controlling not only orienting but also relatively complex vocalizing responses.⁴ In this sense, it could be argued that the conjoining of the speaker and listener repertoires (i.e., bidirectional naming) marks a transition from mutual entailing to combinatorial entailing because the ROE-M involves not only orienting toward an object but also vocalizing the sound that is coordinate with that object. In effect, the object enters a frame of coordination with orienting/evoking *and* vocalizing. Therefore, the relatively complex transformations of function involved in vocalizing responses may be seen as facilitating a shift from mutual to combinatorial entailing. In this sense, it seems appropriate to consider bidirectional naming as a very basic type of relational frame (Greer & Keohane, 2005; Hayes, 1996; Miguel & Petursdottir, 2009).

Furthermore, once the generic response unit of AARR (i.e., the ROE-M) is established, it allows for the evolution of increasingly complex relational responding inside the ROE-M, such as relational networking, the relating of relations (e.g., analogy and metaphor), and the relating of entire relational networks to other relational networks (e.g., extracting common themes from different narratives). This increasing complexity in derived relational responding involves the sophisticated use of symbols and the ability to problem solve in the natural and social environment (see also Miguel, 2018).

Summary

In an updated version of RFT, relatively new concepts have been proposed, such as mutually entailed orienting and evoking, which develop almost in parallel with listening and speaking repertoires. In doing so, the exciting opportunity for connecting and collaborating with colleagues who have been developing VBDT and the concepts of unidirectional, bidirectional, and incidental naming responses becomes increasingly likely and indeed advantageous for all concerned. To move forward with this agenda, it seems useful to consider, if only briefly, all of the points of contact, overlap, and agreement between updated RFT and VBDT and identify, if possible, where they could

work together to form a more comprehensive understanding of the development of complex human behavior.

VBDT AND RFT: SIMILARITIES

VBDT and RFT are two perspectives on human language and its development. In seeking out similarities it is important to recognize, however, that RFT is concerned with both language and cognition, whereas VBDT seems to focus more on verbal behavior or language and on the science of teaching. Of course, RFT has been employed in teaching or educational contexts (see Rehfeldt & Barnes-Holmes, 2009), but it is also very much focused on emotional responses or reactions, given its historical connection to clinical behavior analysis and acceptance and commitment therapy in particular (Hayes et al., 1999). General differences between the two theories may be attributed, in part, to these different applications—that is, RFT's early applications to clinical psychology and VBDT's roots in developmental interventions for children in educational settings. Nonetheless, if we confine our focus to the development of language and BiN, perhaps we will find little to differentiate the two theoretical accounts. As just one example, RFT distinguished between Skinner's (1957) verbal operants (e.g., mands, tacts, autoclitics, and the others) as symbolic versus nonsymbolic—this seems to be one of several aspects where VBDT and RFT intersect. Specifically, RFT suggested that for speaker behavior, for example, for a mand to be symbolic it needed to be part of a relational frame (Barnes-Holmes et al., 2000). In a broadly similar way, VBDT interprets research findings as showing that the cusp for the joining of the listener and speaker needs to be present for the child to be fully verbal. We will elaborate on this argument below.

The importance of early orienting/perceptual processes

Both theories place considerable emphasis, albeit with different terminology, on the behaviors that set the stage for language learning. RFT researchers use the term cooperation (Hayes & Sanford, 2014) as described in evolution science, whereas VBDT researchers refer to preverbal foundational cusps, some of which develop in utero (e.g., demonstrating sensitivity toward mother's voice immediately after birth; DeCasper & Spence, 1986), as being integral to language development (see also Horne & Lowe, 1996, for a discussion on the role of early orienting for naming). In addition, both RFT and VBDT seek to identify specific sequences of experiences or contingencies that lead to the establishment of naming (and other relations). Critically, as outlined above, an updated version of RFT identifies mutually entailed orienting

⁴A reviewer of a previous draft of the paper argued that utterances occurring in the presence of an object may serve as an optimal condition for the object to control future utterances as a discriminative stimulus and that this need not involve a complex derived transformation of function. However, the optimal condition (an utterance in the presence of an object in the first instance) would require a transformation of function, at least as defined above.

and evoking that occur during acts of cooperation as part of a basic step toward AARR including naming. Similarly, VBDT suggests that a developmental trajectory beginning with preverbal behaviors such as orienting to faces and human sounds, followed by producing independent observing responses and speaker responses, results in the acquisition of bidirectional naming.

The importance of the conjoining of two behavioral repertoires in the emergence of bidirectional naming

According to VBDT, as described earlier, the acquisition of incidentally acquired *bidirectional operants* as one of the components of the cusp for the joining of the speaker and the listener is critical to the development of language and a critical marker of social development. In a similar vein, RFT recognizes mutual entailment as one of the core properties of AARR. There are also clear similarities between the two theories in the conceptualization of naming. VBDT researchers describe bidirectional naming as the conjoining of the speaker and listener repertoires and the establishment of a verbal developmental cusp that provides a child with new ways to learn.

According to RFT, derived relational responding always involves a change or modification in the functions of a stimulus in accordance with an entailed relation. For UiN, a relatively limited transformation of functions is involved. The child must only orient toward a novel object (or point toward it/pick it up) after a caregiver names the novel object. For BiN to emerge, the child not only orients toward the object but also utters or vocalizes the sound that was heard when the caregiver named the object. The two accounts (VBDT and RFT) thus appear to be articulating similar functional-analytic ideas using different terms but are essentially in agreement about the behavioral histories involved.

Similar views of incidental BiN

Both theories recognize that bidirectional naming may be considered a type of relational frame (e.g., Greer et al., 2005; Luciano et al., 2007). As noted earlier, for RFT bidirectional naming involves an increase in the complexity of the transformation of functions such that a child is capable of not only orienting toward a named stimulus but also producing the vocal sound that is coordinated with that stimulus. In this sense, RFT and VBDT appear to be in complete agreement. VBDT also emphasizes bidirectional naming as a particularly important cusp that marks the point at which an individual can acquire name-object relations *incidentally*

or in the absence of prior training or direct reinforcement. RFT describes bidirectional naming as the acquisition of a relational frame that is under relatively precise contextual control (in terms of both *Crels* and *Cfuncs*), and thus if these cues are present when a caregiver utters the name of a novel stimulus, then the *incidental* acquisition of that name by the child should be expected. Again, the theories appear to be in general agreement here.

Similar in substance if not in focus

VBDT suggests that it is useful to categorize BiN along a spectrum of behaviors rather than dichotomously measuring its presence or absence. The updated version of RFT, in terms of the HDML framework, includes four dimensions, which allow all instances of AARR to be conceptualized as varying in coherence, complexity, derivation, and flexibility along continua. Thus, the two theories are focused on exploring concepts that are relative rather than absolute. On balance, one of the main differences between VBDT and RFT appears to be the extent to which they have focused empirically on different areas of verbal or language development. For example, VBDT has certainly generated and is generating an increasingly rich data set on BiN to the extent that, paradoxically, RFT researchers could cite more VBDT studies to support the view that naming involves relational framing. Furthermore, VBDT studies and naming research more broadly have provided insights into how basic relational framing may be produced through a young child's interaction with the verbal community.

From the VBDT perspective, BiN has been identified as a critical cusp in allowing for more advanced verbal abilities involved in problem solving and so forth (see Miguel, 2018, for a review of studies showing that participants failed to solve categorization tasks in the absence of listener or speaker naming). It is here that RFT perhaps has generated numerous experimental analyses of human language that extend well beyond BiN. Research on relating relations as the basis for analogical reasoning (see Stewart & Barnes-Holmes, 2004, for a review) and complex relational networks as the basis for rule-governed behavior (e.g., O'Hora et al., 2004; O'Hora et al., 2014) provide good examples. Thus, there is a literature that VBDT researchers could draw on in developing the theory in terms of exploring how BiN feeds into other perhaps important verbal cusps, such as the relating of relations in the development of analogical reasoning in young children (see also Meyer et al., 2019). Once again, recognizing these similarities in substance, if not historical focus (in terms of empirical research areas), appear to provide opportunities and motivation for genuine collaboration now and in future years.

In summary, RFT and VBDT are two theories of language development that have several similarities but

currently exist largely in “parallel universes.” Both theories seek to identify frameworks and empirical research strategies to study behavioral phenomena such as equivalence, naming, and problem solving. Although each theory uses its own terminology, these terms often do not appear to differ fundamentally in the types of functional analyses that they appear to generate (at least at the current time). In the next section, we will briefly consider some examples of how focusing on the overlap between RFT and VBDT may serve to generate new and informative lines of research.

Areas for future research

One area in which a collaborative focus for RFT and VBDT could be of benefit is in the experimental analyses of BiN. Indeed, a recent study provides what we believe to be a strong example (Sivaraman et al., 2021). The authors of this study pointed out that most behavioral research on naming involved presenting an object and its name simultaneously during both training and testing, and thus the training component may establish a transformation of function *directly* between the object and the name. Consequently, successful tests for listener naming may not require the emergence of a novel (entailed) transformation of function. In the study reported by Sivaraman et al. (2021), the researchers presented the object and the name sequentially and nonsimultaneously. They presented the object, and once the child made visual contact, they hid the object from view and then uttered its name. All participants failed to emit listener or speaker responses during the first test session. Four participants then received MET during which they were trained to emit listener responses to two sets with two novel stimuli in each set. All participants were subsequently tested on their responses with a novel object–name relation. Participants who received MET showed improvements in listener behavior using the nonsimultaneous presentation procedure, and one participant showed improvements in speaker behavior.

When an object and its name are presented contemporaneously, then a child “sees object–hears name” and “hears name–sees object” simultaneously. Therefore, both stimulus relations have been directly established, and thus a derived or entailed transformation of function need not be invoked to explain either relational response (e.g., a child does not have to derive name–object from object–name). In contrast, when an object and its name are presented *nonsimultaneously* during training and then speaker and/or listener behavior are observed during test trials, then at the very least the transformation of functions involves an entailed relational response, as defined within RFT. Critically, approaching the research in this manner encourages us to be precise in defining the behavioral processes that may be involved when children learn to name objects in both experimental and natural contexts. The authors of the study considered the RFT concept of BiN as derived transformation of functions and the VBDT

definition as the conjoining of the listener and speaker repertoires to help identify critical variables that generate or fail to generate listener behavior. Future studies in this research vein could expand the analyses to determine the variables that generate *speaker* naming using a nonsimultaneous presentation technique and the effect of longer delays between the presentation of an object and its name.

Another area in which RFT and VBDT could usefully connect with each other involves drawing on the HDML framework and its dimensions (i.e., coherence, complexity, derivation, and flexibility) in attempting to develop experimental analyses of BiN and perhaps other verbal developmental cusps. One example of a future study that uses these dimensions for the study of BiN could address research questions such as whether strengthening the *complexity* of unidirectional naming by establishing contextual control would facilitate improvements in bidirectional naming without specific training. Such a study would involve first training participants to exhibit listener responses (e.g., “point to aardvark”) and subsequently training them to exhibit complexity in those responses (e.g., “point to something that is NOT aardvark”). Probes for bidirectional naming could be conducted after each training phase to determine if increasing complexity by establishing contextual control would increase the likelihood of children showing BiN. Although this is entirely speculative, identifying points of contact between the two theories appears to highlight variables that could be manipulated in studying naming as AARR rather than working from the perspective of one of the theories. In any case, in conducting this type of research, it would seem critical to employ young infants as participants because only then will it be possible to study precise behavioral histories that give rise to various types of naming.

At this point it should be recognized that focusing on connections between VBDT and RFT does not preclude also drawing on BiN research generally. For example, although the focus of VBDT researchers has traditionally been on the identification and establishment of Inc-BiN, the relation between Inc-BiN and other derived relations has become increasingly apparent in the light of recent research. There are multiple previously published empirical studies on this topic in the literature (e.g., Jennings & Miguel, 2017; Meyer et al., 2019; Petursdottir et al., 2015, 2019), and although the relation between Inc-BiN and AARR is only suggestive at this point it does raise some interesting conceptual questions. Specifically, given the strong correlations between Inc-BiN and AARR (Morgan et al., 2021), there seem to be four possibilities: (1) the emergence of other derived relations is a function of Inc-BiN, (2) Inc-BiN is a function of other types of AARR, (3) some other variable is responsible for both, or (4) AARR and Inc-BiN are the same or part of a whole. In terms of scientific explanatory parsimony, the fourth possibility is most appealing and is perhaps most consistent with both VBDT and RFT. As an aside, arguing for

the fourth possibility does not preclude a potentially important role for the relationship between listener and speaker behaviors in *facilitating* the arbitrary relating of heard and spoken words to objects and events in the world. Indeed, this view would be consistent with Inc-BiN and AARR being seen as functionally synonymous.

CONCLUSION

We believe that behavior analysts can contribute a great deal to some of the questions and issues surrounding the evolution of language in humans and the contingencies that operate during its development. Although the arguments presented in the current article are quite (and perhaps sometimes wildly) speculative, it seems like a productive way forward for VBDT and RFT researchers. Identifying the commonalities between VBDT and an updated RFT offers prospects for collaborative research on linguistic phenomena such as naming and the conditions that result in its emergence within a behavior-analytic framework. In addition, the present analysis on mutually entailed orienting and evoking suggests that the historical context for AARR does not begin with listening and speaking but with a basic human cooperative act. This could have implications for empirical and conceptual research on the precursors of naming and the conditions that give rise to mutual entailment and AARR more generally in infants. We hope that the points of mutual interest and contact between researchers outlined in the present analysis result in fruitful collaborations and contributions to the research path pioneered by Sidman (1994) and Skinner (1957).

FUNDING STATEMENT

This study was partially supported by the Marguerite-Marie Delacroix Support Fund awarded to M. Sivaraman and Odysseus Type 1 funding awarded to D. Barnes-Holmes.

ACKNOWLEDGMENTS

The authors would like to thank the anonymous reviewers for their helpful comments on previous drafts of this manuscript.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

ETHICS APPROVAL STATEMENT

No ethical approval was needed for the study. The study was entirely theoretical and did not involve any participants or data collection.

REFERENCES

- Barnes-Holmes, D., Barnes-Holmes, Y., & Cullinan, V. (2000). Relational frame theory and Skinner's *Verbal Behavior*: A possible synthesis. *The Behavior Analyst*, 23(1), 69–84. <https://doi.org/10.1007/BF03392000>
- Barnes-Holmes, D., Barnes-Holmes, Y., Luciano, C., & McEnteggart, C. (2017). From the IRAP and REC model to a multi-dimensional multi-level framework for analyzing the dynamics of arbitrarily applicable relational responding. *Journal of Contextual Behavioral Science*, 6(4), 434–445. <https://doi.org/10.1016/j.jcbs.2017.08.001>
- Barnes-Holmes, D., Barnes-Holmes, Y., & McEnteggart, C. (2020). Updating RFT (more field than frame) and its implications for process-based therapy. *Psychological Record*, 70(4), 605–624. <https://doi.org/10.1007/s40732-019-00372-3>
- Barnes-Holmes, D., Finn, M., McEnteggart, C., & Barnes-Holmes, Y. (2018). Derived stimulus relations and their role in a behavior-analytic account of human language and cognition. *Perspectives on Behavior Science*, 41(1), 155–173. <https://doi.org/10.1007/s40614-017-0124-7>
- Barnes-Holmes, D., & Harte, C. (2022). The IRAP as a measure of implicit cognition: A case of Frankenstein's Monster. *Perspectives on Behavior Science*, 45(3), 559–578. <https://doi.org/10.1007/s40614-022-00352-z>
- Barnes, D., McCullagh, P. D., & Keenan, M. (1990). Equivalence class formation in non-hearing impaired children and hearing impaired children. *Analysis of Verbal Behavior*, 8, 19–30. <https://doi.org/10.1007/BF03392844>
- Barrett, B., Beck, R., Binder, C., Cook, D. A., Engelmann, S., Greer, R. D., Kryklund, S. J., Johnson, K. R., Maloney, M., McCorkle, N., Vargas, J. S., & Watkins, C. L. (1991). The right to effective education. *The Behavior Analyst*, 14(1), 79–82.
- Cahill, C. S., & Greer, R. D. (2014). Actions vs. words: How we can learn both. *Acta Investigacion Psychologica*, 4(3), 1716–1745.
- Cao, Y., & Greer, R. D. (2019). Mastery of echoics in Chinese establishes bidirectional naming in Chinese for preschoolers with naming in English. *The Analysis of Verbal Behavior*, 34(1–2), 79–99. <https://doi.org/10.1007/s40616-018-0106-1>
- Carpenter, M., Nagell, K., & Tomasello, M. (1998). Social cognition, joint attention, and communicative competence from 9 to 15 months of age. *Monographs of the Society for Research in Child Development*, 63(4), 1–174. <https://doi.org/10.2307/1166214>
- Carr, D., Wilkinson, K. M., Blackman, D., & McIlvane, W. J. (2000). Equivalence classes in individuals with minimal verbal repertoires. *Journal of the Experimental Analysis of Behavior*, 74(1), 101–114. <https://doi.org/10.1901/jeab.2000.74-101>
- Choi, J., Greer, R. D., & Keohane, D.-D. (2015). The effects of an auditory match-to-sample procedure on listener literacy and echoic responses. *Behavioral Development Bulletin*, 20(2), 186–206. <https://doi.org/10.1037/h0101313>
- DeCasper, A. J., & Spence, M. J. (1986). Prenatal maternal speech influences newborns' perception of speech sounds. *Infant Behavior & Development*, 9(2), 133–150. [https://doi.org/10.1016/0163-6383\(86\)90025-1](https://doi.org/10.1016/0163-6383(86)90025-1)
- Donley, C. R., & Greer, R. D. (1993). Setting events controlling social verbal exchanges between students with developmental delays. *Journal of Behavioral Education*, 3(4), 387–401.
- 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(1), 5–22. <https://doi.org/10.1901/jeab.2000.73-5>
- Eby, C. M., & Greer, R. D. (2017). Effects of social reinforcement on the emission of tacts by preschoolers. *Behavioral Development Bulletin*, 22(1), 23–24. <https://doi.org/10.1037/bdb0000043>
- Feinman, S. (1980). Infant response to race, size, proximity, and movement of strangers. *Infant Behavior & Development*, 3(3), 187–204. [https://doi.org/10.1016/S0163-6383\(80\)80025-7](https://doi.org/10.1016/S0163-6383(80)80025-7)
- Fienup, D. M. (2018). The future of verbal behavior: Integration. *Analysis of Verbal Behavior*, 34(1–2), 18–23. <https://doi.org/10.1007/s40616-018-0108-z>
- Fienup, D. M., & Brodsky, J. (2020). Equivalence-based instruction: Designing instruction using stimulus equivalence. In M. Fryling, R. A. Rehfeldt, J. Tarbox, & L. J. Hayes (Eds.), *Applied behavior*

analysis of language and cognition: Core concepts and principles for practitioners (pp. 157–173). New Harbinger.

- Finn, M., Barnes-Holmes, D., & McEnteggart, C. (2018). Exploring the single-trial-type-dominance-effect on the IRAP: Developing a differential arbitrarily applicable relational responding effects (DAARRE) model. *The Psychological Record*, *68*(1), 11–25. <https://doi.org/10.1007/s40732-017-0262-z>
- Fiorile, C. A., & Greer, R. D. (2007). The induction of naming in children with no prior tact responses as a function of multiple exemplar histories of instruction. *The Analysis of Verbal Behavior*, *23*(1), 71–87. <https://doi.org/10.1007/BF03393048>
- Friedman, L. F. (2020). *Establishment of increased stimulus control for bidirectional naming increased stimulus control for derived relations in 20-40-month-old toddlers* [Unpublished doctoral dissertation]. Columbia University.
- Gilic, L., & Greer, R. D. (2011). Establishing naming in typically developing two-year-old children as a function of multiple exemplar speaker and listener experiences. *Analysis of Verbal Behavior*, *27*(1), 157–177. <https://doi.org/10.1007/BF03393099>
- Greer, R. D. (1991). Teaching practice to save America's schools: The legacy of B. F. Skinner. *Journal of Behavioral Education*, *1*(2), 159–164.
- Greer, R. D. (2002). *Designing teaching strategies: An applied behavior analysis systems approach*. Academic Press.
- Greer, R. D., Chavez-Brown, M., Nirgudkar, A. S., Stolfi, L., & Rivera-Valdes, C. (2005). Acquisition of fluent listener responses and the educational advancement of young children with autism and severe language delays. *European Journal of Behavior Analysis*, *6*(2), 88–126.
- Greer, R. D., Corwin, A., & Buttigieg, S. (2011). The effects of the verbal developmental capability of bidirectional naming on how children can be taught. *Acta de Investigación Psicológica*, *1*(1), 23–54.
- Greer, R. D., & Du, L. (2015). Experience and the onset of the capability to learn the names of things by exclusion. *The Psychological Record*, *65*(2), 355–373. <https://doi.org/10.1007/s40732-014-0111-2>
- Greer, R. D., & Keohane, D.-D. (2005). The evolution of verbal behavior in children. *Behavioral Development Bulletin*, *12*(1), 31–47. <https://doi.org/10.1037/h0100559>
- Greer, R. D., & Longano, J. (2010). A rose by naming: How we may learn how to do it. *The Analysis of Verbal Behavior*, *26*(1), 73–106. <https://doi.org/10.1007/BF03393085>
- Greer, R. D., Pistoljevic, N., Cahill, C., & Du, L. (2011). Effects of conditioning voices as reinforcers for listener responses on rate of learning, awareness, and preferences for listening to stories in preschoolers with autism. *The Analysis of Verbal Behavior*, *27*(1), 104–124. <https://doi.org/10.1007/BF03393095>
- Greer, R. D., Pohl, P., Du, L., & Moschella, J. L. (2017). The separate development of children's listener and speaker behavior and the intercept as behavioral metamorphosis. *Journal of Behavioral and Brain Science*, *7*(13), 674–704. <https://doi.org/10.4236/jbbs.2017>
- Greer, R. D., & Ross, D. E. (2008). *Verbal behavior analysis: Inducing and expanding complex communication in children with severe language delays*. Allyn & Bacon.
- Hare, B., & Tomasello, M. (2004). Chimpanzees are more skillful in competitive than cooperative cognitive tasks. *Animal Behaviour*, *68*(3), 571–581. <https://doi.org/10.1016/j.anbehav.2003.11.011>
- Harms, G. (2020). *The role of joint attention in the verbal behavior development trajectory* [Unpublished doctoral dissertation]. Columbia University.
- Hart, B. M., & Risley, T. R. (1995). *Meaningful differences in the everyday life of America's children*. Paul Brookes.
- Hart, B., & Risley, T. R. (1999). *The social world of children learning to talk*. Paul Brookes.
- Harte, C., Barnes-Holmes, Y., Barnes-Holmes, D., & McEnteggart, C. (2017). Persistent rule-following in the face of reversed reinforcement contingencies: The differential impact of direct versus derived rules. *Behavior Modification*, *41*(6), 743–763. <https://doi.org/10.1177/0145445517715871>
- Harte, C., Barnes-Holmes, D., Barnes-Holmes, Y., & McEnteggart, C. (2018). The impact of high versus low levels of derivation for mutually and combinatorially entailed relations on persistent rule-following. *Behavioural Processes*, *157*, 36–46. <https://doi.org/10.1016/j.beproc.2018.08.005>
- Hawkins, E., Gautreaux, G., & Chiesa, M. (2018). Deconstructing common bidirectional naming: A proposed classification framework. *The Analysis of Verbal Behavior*, *34*(1–2), 44–61. <https://doi.org/10.1007/s40616-018-0100-7>
- Hawkins, E., Kingsdorf, S., Charnock, J., Szabo, M., & Gautreaux, G. (2009). Effects of multiple exemplar instruction on naming. *European Journal of Behavior Analysis*, *10*(2), 265–273.
- Hayes, S. C. (1991). A relational control theory of stimulus equivalence. In L. J. Hayes & P. N. Chase (Eds.), *Dialogues on verbal behavior* (pp. 19–40). Context Press.
- Hayes, S. C. (1996). Developing a theory of derived stimulus relations. *Journal of the Experimental Analysis of Behavior*, *65*(1), 309–311. <https://doi.org/10.1901/jeab.1996.65-309>
- Hayes, S. C., Barnes-Holmes, D., & Roche, B. (Eds.). (2001). *Relational frame theory: A post-Skinnerian account of human language and cognition*. Plenum Press.
- Hayes, S. C., & Hayes, L. J. (1989). The verbal action of the listener as a basis for rule-governance. In S. C. Hayes (Ed.), *Rule-governed behavior: Cognition, contingencies, and instructional control* (pp. 153–190). Plenum.
- Hayes, S. C., & Sanford, B. T. (2014). Cooperation came first: Evolution and human cognition. *Journal of the Experimental Analysis of Behavior*, *101*(1), 112–129. <https://doi.org/10.1002/jeab.64>
- Hayes, S. C., Strosahl, K. D., & Wilson, K. G. (1999). *Acceptance and commitment therapy: An experiential approach to behavior change*. Guilford Press.
- Horne, P. J., Hughes, J. C., & Lowe, C. F. (2006). Naming and categorization in young children: IV: Listener behavior training and transfer of function. *Journal of the Experimental Analysis of Behavior*, *85*(2), 247–273. <https://doi.org/10.1901/jeab.2006.125-04>
- Horne, P. J., & Lowe, C. F. (1996). On the origins of naming and other symbolic behavior. *Journal of the Experimental Analysis of Behavior*, *65*(1), 185–241. <https://doi.org/10.1901/jeab.1996.65-185>
- Horne, J., Lowe, C. F., & Randle, V. R. L. (2004). Naming and categorization in young children II: Listener behavior training. *Journal of the Experimental Analysis of Behavior*, *81*(3), 267–288.
- Hotchkiss, R. M., & Fienup, D. M. (2019). A parametric analysis of a protocol to induce bidirectional naming: Effects of protocol intensity. *The Psychological Record*, *70*(3), 481–497. <https://doi.org/10.1007/s40732-020-00383-5>
- Hranchuk, K., Greer, R. D., & Longano, J. (2019). Instructional demonstrations are more efficient than consequences alone for children with naming. *Analysis of Verbal Behavior*, *35*(1), 1–20. <https://doi.org/10.1007/s40616-018-0095-0>
- Jennings, A. M., & Miguel, C. F. (2017). Teaching intraverbal bidirectional naming to establish generalized equivalence class performances. *Journal of the Experimental Analysis of Behavior*, *108*(2), 269–289.
- LaFrance, D. L., & Tarbox, J. (2019). The importance of multiple exemplar instruction in the establishment of novel verbal behavior. *Journal of Applied Behavior Analysis*, *53*(1), 10–24.
- Lipkens, R., Hayes, S. C., & Hayes, L. J. (1993). Longitudinal study of the development of derived relations in an infant. *Journal of Experimental Child Psychology*, *56*(2), 201–239. <https://doi.org/10.1006/jecp.1993.1032>

- Lodhi, S., & Greer, R. D. (1989). The speaker as listener. *Journal of the Experimental Analysis of Behavior*, 51(3), 353–359. <https://doi.org/10.1901/jeab.1989.51-353>
- Longano, J., & Greer, R. D. (2014). Is the source of reinforcement for naming multiple conditioned reinforcers for observing responses? *The Analysis of Verbal Behavior*, 31(1), 96–117. <https://doi.org/10.1007/s406160022-y>
- Lowe, C. F., Horne, P. J., & Hughes, J. C. (2005). Naming and categorization in young children: III Vocal tact training and transfer of function. *Journal of Experimental Analysis of Behavior*, 83(1), 47–65.
- Lowe, C. F., Horne, P. J., Harris, D. S., & Randle, V. R. L. (2002). Naming and categorization in young children: Vocal tact training. *Journal of the Experimental Analysis of Behavior*, 78(3), 527–549.
- Lowenkron, B. (1991). Joint control and the generalization of selection-based verbal behavior. *The Analysis of Verbal Behavior*, 9, 121–126.
- Lowenkron, B. 1998. Some logical functions of joint control. *Journal of the Experimental Analysis of Behavior*, 69(3), 327–354.
- 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 the Experimental Analysis of Behavior*, 87(3), 349–365. <https://doi.org/10.1901/jeab.2007.08-06>
- Luciano, M. C., Herruzo, J., & Barnes-Holmes, D. (2001). Generalization of say-do correspondence. *The Psychological Record*, 51(1), 111–130.
- Luciano, M. C., & Polaino-Lorente, A. (1986). Effects of the acquisition of prerequisite behavior on the learning of nonvocal verbal behavior and vocal imitation in children with severe retardation. *Psychological Record*, 36(3), 315–332. <https://doi.org/10.1007/BF03394951>
- Ma, M. L., Miguel, C. F., & Jennings, A. M. (2016). Training intraverbal naming to establish equivalence class performances. *Journal of the Experimental Analysis of Behavior*, 105(3), 409–426. <https://doi.org/10.1002/jeab.203>
- Maffei, J., Singer-Dudek, S., & Keohane, D. (2014). The effects of the establish of adult faces and/or voices as conditioned reinforcers for children with ASD and related disorders. *Acta de Investigación Psicológica*, 4(3), 1621–1641. [https://doi.org/10.1016/S2007-4719\(14\)70970-6](https://doi.org/10.1016/S2007-4719(14)70970-6)
- Miguel, C. F. (2016). Common and intraverbal bidirectional naming. *The Analysis of Verbal Behavior*, 32(2), 125–138.
- Miguel, C. F. (2018). Problem-solving, bidirectional naming, and the development of verbal repertoires. *Behavior Analysis: Research and Practice*, 18(4), 1–14. <https://doi.org/10.1037/bar0000110>
- Miguel, C. F., & Petursdottir, A. I. (2009). Naming and frames of coordination. In R. A. Rehfeldt & Y. Barnes-Holmes (Eds.), *Derived relational responding: Applications for learners with autism and other developmental disabilities* (pp. 129–148). New Harbinger.
- Meyer, C. S., Cordeiro, M. C., & Miguel, C. F. (2019). The effects of listener training on the development of analogical reasoning. *Journal of the Experimental Analysis of Behavior*, 112(2), 144–166. <https://doi.org/10.1002/jeab.549>
- Morgan, G., Greer, R. D., & Fienup, D. M. (2021). Descriptive analyses of the relations between bidirectional naming, arbitrary, and non-arbitrary relations. *The Psychological Record*, 71(3), 367–387. <https://doi.org/10.1007/s49732-020-00408-z>
- Mundy, P., Block, J., Vaughan Van Hecke, A., Delgado, C., Parlade, M., & Pomeroy, Y. (2007). Individual differences in the development of joint attention in infancy. *Child Development*, 78(3), 938–954. <https://doi.org/10.1111/j.1467-8624.2007.01042.x>
- Novak, G., & Pelaez, M. (2004). *Child and adolescent development: A behavioral systems approach*. Sage Publications.
- 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(3), 437–460. <https://doi.org/10.1007/BF03395484>
- O’Hora, D., Barnes-Holmes, D., & Stewart, I. (2014). Antecedent and consequential control of derived instruction-following. *Journal of the Experimental Analysis of Behavior*, 102(1), 66–85. <https://doi.org/10.1002/jeab.95>
- Olaff, H. S., & Holth, P. (2020). The emergence of bidirectional naming through sequential operant instruction following the establishment of conditioned social reinforcers. *Analysis of Verbal Behavior*, 36(1), 21–48. <https://doi.org/10.1007/s40616-019-00122-0>
- Paniagua, F. A., & Baer, D. M. (1982). The analysis of correspondence training as a chain reinforceable at any point. *Child Development*, 53(3), 786–98.
- Pelaez, M., Virues-Ortega, J., & Gewirtz, J. L. (2012). Acquisition of social referencing via discrimination training in infants. *Journal of Applied Behavior Analysis*, 45(1), 23–36. <https://doi.org/10.1901/jaba.2012.45-23>
- Pérez-González, L. A., Cereijo-Blanco, N., & Carnerero, J. J. (2014). Emerging tacts and selections from previous learned skills: A comparison between two types of naming. *The Analysis of Verbal Behavior*, 30(2), 184–192. <https://doi.org/10.1007/s40616-014-0011-1>
- Petursdottir, A. I., Carp, C. L., Peterson, S. P., & Lepper, T. L. (2015). Emergence of visual–visual conditional discriminations. *Journal of the Experimental Analysis of Behavior*, 103(2), 332–348. <https://doi.org/10.1002/jeab.136>
- Petursdottir, A. I., Cox, R. E., McKeon, C. A., & Mellor, J. R. (2019). Baseline training sequence affects speed of emergent conditional discriminations. *Journal of the Experimental Analysis of Behavior*, 112(1), 60–73. <https://doi.org/10.1002/jeab.53>
- Pohl, P., Greer, D. R., Du, L., & Moschella, L. J. (2018). Verbal development, behavioral metamorphosis, and the evolution of language. *Perspectives on Behavior Science*, 43(1), 215–232. <https://doi.org/10.1007/s40614-018-00180-0>
- Rehfeldt, R. A., & Barnes-Holmes, Y. (2009). *Derived relational responding: Applications for learners with autism and other developmental disabilities*. New Harbinger Publications.
- Rosales-Ruiz, J., & Baer, D. M. (1996). A behavior-analytic view of development. In S. Bijou & E. Ribes (Eds.), *New directions in behavior development* (pp. 155–180). Context Press.
- Rosales-Ruiz, J., & Baer, D. M. (1997). Behavioral cusps: A developmental and pragmatic concept for behavior analysis. *Journal of Applied Behavior Analysis*, 30(3), 533–544.
- Schmelzkopf, J., Greer, R. D., Singer-Dudek, J., & Du, L. (2017). Experiences that establish interest in speaking and listening to others. *Behavioral Development Bulletin*, 22(1), 44–66. <https://doi.org/10.1037/bdb0000026>
- Sidman, M. (1971). Reading and auditory-visual equivalences. *Journal of Speech and Hearing Research*, 14(1), 5–13.
- Sidman, M. (1994). *Equivalence relations and behavior: A research story*. Authors Cooperative.
- Sidman M. (2000). Equivalence relations and the reinforcement contingency. *Journal of the Experimental Analysis of Behavior*, 74(1), 127–146. <https://doi.org/10.1901/jeab.2000.74-127>
- Sidman, M., Rauzin, R., Lazar, R., Cunningham, S., Tailby, W., & Carrigan, P. (1982). A search for symmetry in the conditional discriminations of rhesus monkeys, baboons, and children. *Journal of the Experimental Analysis of Behavior*, 37(1), 23–44. <https://doi.org/10.1901/jeab.1982.37-23>
- Sivaraman, M., Barnes-Holmes, D., & Roeyers, H. (2021). Nonsimultaneous stimulus presentations and their role in listener naming. *Journal of the Experimental Analysis of Behavior*, 116(3), 300–313. <https://doi.org/10.1002/jeab.715>
- Sivaraman, M., Virues-Ortega, J., Maes-Dhaeyere, L., Saelaert, E., & Roeyers, H. (2022). Social referencing training in children with autism spectrum disorder: A randomized controlled study. *Journal*

- of *Applied Behavior Analysis*, 55(3), 763–781. <https://doi.org/10.1002/jaba.935>
- Skinner, B. F. (1954). *Science and human behavior*. Appleton-Century-Crofts.
- Skinner, B. F. (1957). *Verbal behavior*. Copley Publishing Group and the B. F. Skinner Foundation.
- Slaughter, V., & McConnell, D. (2003). Emergence of joint attention: Relationships between gaze following, social referencing, imitation, and naming in infancy. *Journal of Genetic Psychology*, 164(1), 54–71. <https://doi.org/10.1080/00221320309597503>
- Stewart, I., & Barnes-Holmes, D. (2004). Relational frame theory and analogical reasoning: Empirical investigations. *International Journal of Psychology and Psychological Therapy*, 4(2), 241–262.
- Tenenbaum, E. J., Sobel, D. M., Sheinkopf, S. J., Malle, B. F., & Morgan, J. L. (2014). Attention to the mouth and gaze following in infancy predict language development. *Journal of Child Language*, 41(1), 1–17. <https://doi.org/10.1017/S0305000914000725>
- Tomasello, M. (1988). The role of joint attentional process in early language development. *Language Sciences*, 10(1), 69–88.
- Tomasello, M., & Todd, J. (1983). Joint attention and lexical acquisition style. *First Language*, 4(12), 197–211. <https://doi.org/10.1177/014272378300401202>
- Williams, G., & Greer, R. D. (1993). A comparison of verbal-behavior and linguistic communication curricula for training developmentally delayed adolescents to acquire and maintain vocal speech. *Behaviorology*, 1, 31–46.
- Wilson, D. S., Hayes, S. C., Biglan, A., & Embry, D. D. (2014). Evolving the future: Toward a science of intentional change. *The Behavioral and Brain Sciences*, 37(4), 395–416. <https://doi.org/10.1017/S0140525X13001593>

How to cite this article: Sivaraman, M., Barnes-Holmes, D., Greer, R. D., Fienup, D. M., & Roeyers, H. (2023). Verbal behavior development theory and relational frame theory: Reflecting on similarities and differences. *Journal of the Experimental Analysis of Behavior*, 1–15. <https://doi.org/10.1002/jeab.836>