Relational frame theory 20 years on: The Odysseus voyage and beyond

Published in:
Journal of the Experimental Analysis of Behavior

Publication Status:
Published (in print/issue): 15/03/2022

DOI:
10.1002/jeab.733

Document Version
Publisher's PDF, also known as Version of record

Journal of the
Experimental Analysis of Behavior

Journal of the Experimental Analysis of Behavior 2022, 1–27

Relational frame theory 20 years on: The Odysseus voyage and beyond

Dermot Barnes-Holmes[1] and Colin Harte[2,3]

[1]School of Psychology, Ulster University, Northern Ireland

[2]Departmento de Psicologia, Universidade Federal de São Carlos, Brazil

[3]Paradigma – Centro de Ciências e Technologia do Comportamento, Brazil

The seminal text on relational frame theory (RFT) was published 20 years ago and purported to offer a single overarching behavior-analytic account of human language and cognition. In the years thereafter, an increasing number of empirical and conceptual articles, book chapters in edited volumes, and whole volumes devoted to the account emerged. In recent years, RFT has experienced a period of intense empirical and conceptual development, facilitated in part by a research grant awarded by the Flanders Science Foundation, under its Odysseus program. This research program aimed to advance and extend the RFT account beyond the rendition presented in the seminal Hayes et al. (2001) volume. The current article aims to provide an overview of this research program, the empirical work and concepts it gave rise to, and their implications for an RFT account of human symbolic language and cognition. Overall, therefore, the article provides an account of relatively recent developments in RFT that extend beyond the 2001 volume and thus will, we hope, inform future research and critiques of the theory going forward.

The seminal text on relational frame theory (RFT) was published 20 years ago (Hayes et al., 2001). The volume purported to offer a single overarching behavior-analytic account of human language and cognition. The first public presentation of the core RFT idea, that relating is an operant, was made 16 years earlier at the Association for Behavior Analysis meeting in Columbus, Ohio entitled "Verbal behavior, equivalence classes, and rules: New definitions, data, and directions" (Hayes & Brownstein, 1985). More elaborated (early) versions of RFT emerged as chapters in a volume on rule-governed behavior (Hayes & Hayes, 1989) and in another book, "*Dialogues on verbal behavior*," (Hayes, 1991), which was based on an invited address from a conference held in 1986 (Hayes, 1986). Numerous journal articles appeared in various behavior-analytic journals in the late 1980s (e.g., Devaney et al., 1986) and throughout the 1990s (e.g., Barnes & Holmes, 1991; Steele & Hayes, 1991; Dymond & Barnes, 1995) that either referred directly to the theory or presented studies designed to test some aspect of the account.

Following publication of the full book-length treatment of RFT (Hayes et al., 2001) a growing number of empirical and conceptual articles and book chapters began to emerge, with edited volumes, or part thereof, devoted to the theory (Dymond & Roche, 2012; Zettle

The bulk of the material presented in the current article, both conceptual and empirical, was greatly facilitated by a large research grant awarded by the Flanders Science Foundation, under its prestigious Odysseus programme, which provides funding to establish a research team for 5 years at a Flemish university, in this case Ghent University. Although the current article was written by two individuals from that programme following the completion of the Odysseus research project, it is important to acknowledge the important contribution of all members of the Odysseus research team to the conceptual and empirical developments presented herein, listed here alphabetically: Yvonne Barnes-Holmes, Martin Finn, Deirdre Kavanagh, Aileen Leech, Ciara McEnteggart, Michel Quak, and Roberta Vastano. We would also like to express our gratitude to Professors Jan De Houwer, Geertz Crombez, Rudi De Raedt, and many of the other researchers and support staff in the Faculty of Psychology and Educational Sciences at Ghent University, who provided so much support, help and assistance to the Odysseus research team during their 5 year "voyage" in Ghent. The authors also wish to acknowledge the important contribution that the on-going collaboration with Maithri Siveraman has had on some of the material presented in the current article, particularly in relation to the concept of orienting in the development of relational responding in early childhood. Preparation of the current manuscript was supported by a research fellowship awarded to the second author by the São Paulo Research Foundation (FAPESP, Grant #2019/24210-0).

Address correspondence to: Dermot Barnes-Holmes, School of Psychology, Ulster University, Coleraine, Northern Ireland, UK. Email: d.barnes-holmes@ulster.ac.uk

doi: 10.1002/jeab.733

et al., 2016). Relatively recent reviews provided extensive summaries of the research that emerged on RFT before and primarily following publication of the 2001 volume (Hughes & Barnes-Holmes, 2016a, 2016b). A year later an article appeared in the *Journal of Contextual Behavioral Science* that marked the beginning of a period of intense "up-dating" or conceptual development of RFT that aimed to advance or extend beyond the rendition of the theory as presented in the 2001 volume (Barnes-Holmes et al., 2017).

The current article will focus on the research activities and conceptual developments of a single research group (see acknowledgements), and their collaborators, that emerged during the "Odysseus voyage"[1] (2015-2020) and from that voyage to the present time. The current article is thus very much focused on recent conceptual and empirical developments that emerged from a specific research program, rather than a scholarly historical narrative on recent RFT research in general and/or the various ways in which the theory has been used to inform or drive behavior-analytic applied research and practice (e.g., Dixon, 2016; Villatte et al., 2016). More importantly, the current article will provide readers with an account of RFT that clearly extends beyond the 2001 volume and thus should be taken on board in critiquing the theory, either positively or negatively.

In focusing on the contribution of a single research group and their collaborators for the current article, it is important to be absolutely clear that what we offer here is just one recent attempt to extend and develop RFT as articulated in the seminal volume (Hayes et al. 2001). We are not, therefore, arguing, or even suggesting, that the new terms and concepts presented herein constitute the "new" RFT account. Some of the ideas we offer may take hold within the wider research community and others may not—only time will tell. Indeed, even the seminal RFT volume contained concepts that have rarely been the direct target of experimental analyses of behavior in the RFT literature (e.g., pragmatic verbal analysis, pliance, tracking, and augmenting; see Harte & Barnes-Holmes, 2021a, for an extended discussion). As was the case for the seminal volume, therefore, we are simply attempting to provide analytic concepts that the wider community may find useful. Before moving on to the presentation of the new terms and concepts that emerged from the Odysseus research program, and beyond, we will briefly cover the emergence of RFT in behavior analysis.

[1]The bulk of the material presented in the current article, both conceptual and empirical, was greatly facilitated by a large research grant awarded by the Flanders Science Foundation under its prestigious Odysseus program, which provides funding to establish a research team for 5 years at a Flemish University, in this case Ghent University.

The Emergence of RFT within Behavior Analysis

Fifty years ago, a major figure in behavior analysis, Murray Sidman (1971), reported an effect that appeared to be closely associated with human language. Specifically, Sidman was exploring ways through which to teach basic reading abilities to a teenage boy with severe learning disabilities. Sidman and his colleagues had first taught the individual to match spoken words to both pictures and printed words. Following this training, the individual spontaneously matched the printed words to the pictures, and vice versa, without direct reinforcement. In effect, upon reinforcing a subset of reading (relational) responses, a number of unreinforced matching responses emerged. This general phenomenon and its study subsequently became known as the analysis of stimulus equivalence relations (see Sidman, 1994, for a book length overview of the early history of this research program). Critically, although these types of untrained, emergent responses were repeatedly shown in humans with relative ease, they were generally absent in nonhumans (or at best very weak; e.g., Dugdale and Lowe, 2000; Sidman et al., 1982; the reader is also referred to commentaries in Dougher et al., 2014, which suggest that clear evidence of stimulus equivalence, as defined by Sidman, has yet to be observed in nonhuman species).

Two key issues thus emerged from the stimulus equivalence phenomenon: 1) Equivalence was difficult to explain simply in terms of direct contingencies of reinforcement because responses that had not previously been reinforced readily emerged; 2) the phenomenon was discovered while teaching basic reading skills, and nonhumans repeatedly failed to show clear evidence of equivalence responding. Critically, these two points suggested some connection between stimulus equivalence and human language. In working on this apparent connection, a number of theoretical or conceptual positions emerged. One explanation suggested by Sidman et al. (1982) was that equivalence itself may be a basic stimulus function unique to humans and could thus provide a basis for explaining human language (or more precisely, symbolic relations) itself. Other researchers, however, suggested the reverse; that is, some argued that human language (and naming in particular) provided the basis for stimulus equivalence (e.g., Horne & Lowe, 1996).

A third explanation for the emergent properties of equivalence relations was proposed within the account known as RFT. Specifically, it was argued that the relating behavior observed in emergent equivalence responding could itself be considered a class of generalized operant behavior (Hayes, 1991). That is, equivalence responding was learned during early language development, and therefore equivalence and language (or symbolic relations) could be thought of as functionally synonymous. Hayes also argued that equivalence was but one class of generalized operant and that a wide variety were possible, referred to as relational frames. Thus, during early language acquisition, as children are taught to respond in accordance with equivalence relations (or frames of coordination), they are also taught to respond in accordance with many other relational frames such as difference, opposition, comparison, and so on. As such, various patterns of emergent (or derived) relational responses should be possible and thus help to provide a behavior-analytic account of the complexities involved in human language and cognition generally. To fully appreciate this argument, and the extension of RFT beyond basic stimulus equivalence, we will now provide a brief description of the core RFT concepts.

Core Concepts in RFT (and Their Technical Explanation)

RFT posited that there were three fundamental properties involved in the basic operant unit of the relational frame. The properties involved in the relational frame are abstract in a functional sense because they aim to account for the different generalized patterns of derived stimulus relating possible from an RFT perspective (e.g., frames of coordination/equivalence, opposition, comparison, difference, etc.). As a brief example, consider the difference in patterns of derived relating that emerge from coordination (e.g., the relation of "same") versus comparison (e.g., bigger/smaller); specifically, if A is the same as B, then B is the same as A, but if A is bigger than B, then B is smaller than A. Or, in other words, the relation of coordination is symmetrical, but the relation of comparison is not.

The concept of the relational frame (a "minimal" network of relational responses) is defined as a generalized pattern of arbitrarily applicable relational responding (AARR); that is, relational responses that are not based solely on the formal properties of the stimulus relations (e.g., the word "tiny" has four letters and means very small, but the word "big" has three letters and means very large). Relational frames are generally comprised of three properties: (i) mutual entailment, (ii) combinatorial entailment, and (iii) the transformation of stimulus functions. Mutual entailment refers to a bidirectional relation between two stimuli. For example, if X is smaller than Y, this mutually entails that Y is bigger than X. Combinatorial entailment refers to novel relations that emerge between and among stimuli when three or more stimuli are related. For example, if X is more than Y (mutually entailing that Y is less than X) and Y is more than Z (mutually entailing that Z is less than Y) then additional relations will emerge such that X is more than Z and Z is less than X. The transformation of stimulus functions refers to the change or transformation of functions of one or more stimuli in a relational frame that results from a change in functions of other stimuli in that frame. For example, if three stimuli (A, B, and C) participate in a frame of equivalence, and A is paired with an aversive stimulus (e.g., shock), B and C may also acquire aversive (shock) functions. Crucially, the transformation of stimulus functions emerges in the absence of direct reinforcement, instruction, or prompting. This property thus emphasizes the role that symbolic relations in human language play in stimuli losing, gaining, or changing (i.e., transforming) their psychological properties.

The reader should note that the term transformation of function is employed within RFT to emphasize the fact that the functions of stimuli are transformed in accordance with the specific entailment properties of the relations within which they participate (e.g., Dymond &

Barnes, 1995). For example, consider a situation in which a child has been bitten by a friend's small pet white mouse. The child later learns that another friend has recently bought a white rabbit as a pet. Based on the transformation of fear functions, in accordance with the frame of comparison (in this case, smaller/larger), it is possible that the friend's rabbit may evoke even greater fear and avoidance than the smaller mouse that actually bit the child in the first place (see Dougher et al., 2007, for relevant experimental evidence). An explanation based solely on primary stimulus generalization (i.e., whiteness, fluffiness, etc.) could predict some level of fear/avoidance for the rabbit but it should be less fear, not more, based on the many ways in which mice and rabbits differ in physical appearance.

The distinction between the transformation of stimulus functions and relational entailment is considered to be an important one within RFT. That is, doing so distinguishes between the act of abstractly relating stimuli in myriad complex ways from the impact of that relating on the functions of those stimuli. Furthermore, in making this distinction, RFT specifies that these properties (entailment and transformation of functions) are under different classes of contextual control. Specifically, Crel contextual cues control the type of relation (e.g., coordination, comparison, difference, etc.), thus determining the entailment properties, and Cfunc contextual cues control the behavioral functions produced during this relating, thus determining the transformation of function properties. For RFT, therefore, both types of contextual control are crucial in analyzing how entailment and transformations of functions combine in any given instance of AARR. For example, imagine you were in a café in Brazil with a friend and they told you that the word 'bolo' on the menu means cake. The word "means" could function here as a Crel for coordination between the word "bolo" and what you have learned previously to be a cake. If your friend then asks "what's the nicest bolo you have ever eaten," the words "nicest" and "eaten" may serve as Cfuncs that evoke the gustatory properties of a particularly tasty cake you may have eaten at some time. Of course, "nicest" and "eaten" may also be entailed relationally with other events (e.g., the worst cake you have consumed). However, we are highlighting the Cfunc properties of these words in the current example to illustrate how RFT uses both defining properties of a frame (Crel and Cfunc contextual control) to describe how verbal stimuli produce their effects in the natural environment of the wider verbal community.

We have just outlined the core properties of relational framing, the basic behavioral unit for analyzing human language and cognition from a traditional RFT perspective. However, it is important to emphasize that for RFT this unit is considered, in part, a product of learning within the lifetime of the individual. That is, AARR is considered to be learned behavior for which RFT seeks to provide an explanation. As an illustrative example, consider how learning to respond in accordance with the frame of coordination draws on the learning history involved in learning to name, and is thus considered to be one of the most basic classes of AARR. Upon hearing the name of an object, young children often learn to look at or point to that object, along with reinforcement for vocally producing the name of the object. Many such instances of coordinating numerous objects with their respective names across numerous contexts thereby establishes the operant class of coordination. Thus, derived coordination is established for the child, such that direct learning (e.g., involving differential reinforcement) is no longer necessary in the presence of novel objects. For example, imagine the child is then shown a novel object and told its name. The child may now produce the name of the object without further direct training. That is, once the generalized relational response of coordinating objects with their names is established in the child's behavioral repertoire, simply hearing the name of a novel object in the presence of that object may "spontaneously" produce the appropriate naming response. Critically, once this pattern of responding has been effectively established through multiple exemplar training, the generalized relational response may then be applied to any stimuli given appropriate contextual cues (e.g., "that's an aardvark" coordinates the vocal sound with the animal in the absence of any explicit reinforcement).

Beginning to Advance RFT beyond the 2001 Volume

Hayes and Sanford (2014) argued that the human ability to learn to relate stimuli in an arbitrarily applicable manner may have emerged from the evolution of highly cooperative behaviors within the human species. This emphasis was far more explicit than was recognized in the 2001 volume; the role of human cooperation was not ignored in that text but was certainly not emphasized or considered in any great detail. The emphasis on human cooperation as the key driver for the evolution of human language and cognition is generally consistent with a view that is emerging within some of the RFT literature (e.g., Hayes & Sanford, 2014) that AARR is perhaps unique to the human species, or at least relatively advanced or complex examples of AARR constitute unique human behaviors. The uniqueness of AARR is perhaps best reflected in proposing a new generic response unit of analysis within RFT, referred to as the ROE-M (pronounced 'roam'). The acronym stands for relating, orienting, evoking, and motivating, and the basic argument is that analyzing human behavior would benefit by conceptualizing virtually all psychological acts as involving a dynamical and nonlinear interaction among the elements that comprise the ROE-M. We will discuss the ROE-M in detail later but very briefly, *relating* refers to AARR as defined within RFT; *orienting* refers to the extent to which a stimulating event is noticed or "stands out" in the wider context; *evoking* refers to the extent to which a stimulating event is deemed to be appetitive versus aversive; and *motivating* refers to the putative strength of motivational variables, which interact with orienting and/or evoking functions, and indeed relating, in a dynamical manner. To appreciate the core argument we are making for the ROE-M as a new generic unit of analysis for RFT will require that we first describe the broader conceptual developments surrounding the ROE-M, followed by a summary of the relevant empirical findings underpinning these developments. From this point, therefore, the current article will cover primarily the conceptual developments as well as the empirical work that helped generate these developments, and also some very recent empirical research that *was* generated, at least in part, by these developments.

Entailed Orienting, Evoking, and Motivating

The recent RFT focus on the unique evolution of symbolic language and cognition in homo sapiens and related primates has led us to propose that orienting, evoking, and motivating events be labelled entailed orienting, entailed evoking, and entailed motivating when they occur in an historical context that could be seen as involved in linking these behaviors to relating (i.e., AARR), and thus establishing the ROE-M itself. These terms are invoked here simply to recognize that even some behaviors that occur prenatally for members of the human species could be seen as potentially important for the subsequent development of AARR. For example, research has indicated that prenatal humans show sensitivity to the sounds of human voices in the womb, in that they acquire specific "listener" responses more readily for those sounds post-embryonically (e.g., DeCasper & Spence, 1986; Moon et al., 1993; Moon et al., 2013; see Gervain, 2015, for a recent review). Although extremely speculative at this time, in terms of supporting empirical evidence, the example does serve to emphasize that the historical context for the emergence of the ROE-M may be traced back to interactions that occur well before infants are engaged in formal language learning episodes. As another example, evidence indicates that both humans and canines are both biologically or genetically "prepared" to cooperate with other humans, with puppies even "outperforming" other primates in effective human cooperation at the same early stage of development (Bray et al., 2021). Crucially, however, only members of the human species then go on to show strong evidence of AARR, or at least the relatively complex forms of AARR that characterize human symbolic behavior. We have specifically invoked the concepts of *entailed* orienting, evoking, and motivating, therefore, to delineate human behavioral interactions that are seen as central to the development of the ROE-M, but are not yet fully part of it. Thus, an infant's orienting (startle) response to a loud bang, for example, would not be labelled as entailed unless it had some role in preparing the infant to relate such an event to an arbitrary (symbolic) stimulus at some point in the future. In effect, we are referring to the behavioral history that *precedes* the ROE-M, but is seen as typically involved in helping to create this behavioral unit. Subsequently, we label the combination of the three elements (entailed orienting, evoking, and motivating) as the entailment triangle (see below), which may be seen as a relatively strong characteristic of the human species. Adopting this admittedly speculative assumption seems to encourage or facilitate explanatory depth between evolutionary science and a behavior-analytic account of human language and cognition.

As an aside, we recognize and fully acknowledge that defining the behavioral events involved in the entailment triangle in terms of behaviors that emerge only months or even years later involves a high degree of molar theorizing (on balance, such molarity is entirely consistent with RFT). In any case, empirically testing this conceptual analysis would involve demonstrating that some level of AARR, such as contextually controlled generalized symmetry, is functionally related to the behavioral events involved in the entailment triangle.

Before moving on, we should explain that the three elements—orienting, evoking, and motivating—are combined into a unit because it seems useful to avoid separating them. In doing so, we recognize that this is a relatively assumptive strategy, but in conceptualizing psychological units of analysis as behavioral events it seems important to acknowledge that when a stimulus acquires an orienting function, evoking and motivating functions are also involved. For example, when a prenatal child is "sensitized" to the sound of the mother's voice, this could be interpreted as involving orienting, evoking, and motivating functions. That is, when a vocal sound is noticed (i.e., orienting function), and is paired with the delivery of "comforting" nutrients inside the womb (i.e., evoking functions), it could be seen as altering the motivational functions of the mother's voice. We summarize this assumptive analysis with what we call the entailment triangle (see Fig. 1).[2] Again, we recognize that the entailment triangle concept is highly speculative but may be useful because it avoids separating stimulus functions, at the level of process, in a manner that is difficult to maintain when the unit of analysis is a behavioral event. For example, in behavior analysis a stimulus that is defined as a reinforcer will also possess eliciting and motivational functions, even if these properties are not involved in a specific experimental or applied analysis of behavior.

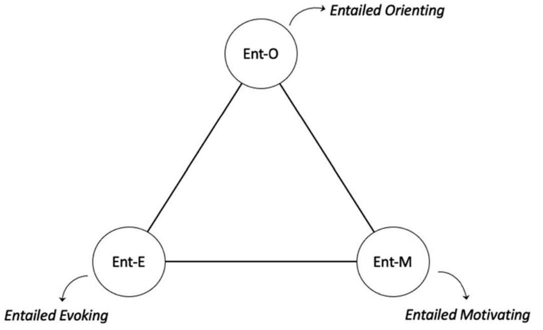
[2]Both the ROE-M and entailment triangle concepts are relatively new within RFT. Until recently, the ROE-M was simply referred to as the ROE (see Barnes-Holmes et al., 2020, 2021), but has since been modified to accommodate the impact of motivating variables in determining psychological events (Harte & Barnes-Holmes, 2021b). The entailment triangle concept emerged during the writing of this very article in attempting to distinguish more clearly between learning at different stages of human development, in and outside of the womb. Indeed, both concepts emerged after the Odysseus research programme was complete, and this is reflected in the "and beyond" phrase contained in the title of the current article.

Mutually Entailed Orienting, Evoking and Motivating

As indicated above, RFT has increasingly focused on cooperation as a key driver of AARR (see Hayes & Sanford, 2014, for a detailed argument). For RFT, a more explicit and prominent cooperation-came-first focus suggests that the critical behavioral history for AARR may not begin with simply speaking or even listening, but rather with what has been termed mutually entailed orienting (Barnes-Holmes & Sivaraman, 2020). Mutually entailed orienting refers to a class of orienting behavior that only occurs in the context of a cooperative act involving a caregiver and a human infant. We have argued elsewhere that a potential marker for mutually entailed orienting is characterized by a child looking back and forth between a caregiver and an object or stimulus that the caregiver has pointed at or named (Harte & Barnes-Holmes, 2021b). As an aside, it may be tempting to define this as an example of joint control, as defined by Lowenkron (1991), but mutually entailed orienting may simply involve orienting back and forth between an object and a person who has pointed at that object without any topographical matching of a mediating (e.g., vocal) response with a response produced by the observed stimulus. Of course, mutually entailed orienting clearly involves "basic" orienting responses, but the critical point here is that they occur as part of a human cooperative act between two or more people. Naturally, orienting per se remains a relatively simple response to any event functioning as a stimulus (i.e., an object or event that is simply noticed). Indeed, a stimulus cannot be defined as a stimulus without some orienting property. Thus, a preverbal infant may show a startle (orienting) response (and start to cry) upon hearing an unexpected loud bang when alone. However, in this instance the infant is not engaging in what we term mutually entailed orienting because it is not part of a

Figure 1

Visual Representation of the Entailment Triangle



Note. The Entailment Triangle aims to recognize that it may be useful to distinguish even prenatal behaviors for humans from topographically similar responses in other species.

cooperative act; indeed, it may not even be considered entailed orienting unless, as noted above, it is involved in preparing the infant, at some point in the future, to acquire symbolic relational responding as defined by RFT.

Mutually entailed orienting is viewed as likely important because it facilitates the establishment of appetitive and aversive evoking functions for stimuli in the child's environment by the caregiver. As soon as an infant and caregiver are engaging in mutually entailed orienting, the infant can be oriented by the caregiver toward particular stimuli and encouraged to avoid stimuli that are considered "dangerous" or approach stimuli that are considered "safe." Thus, while mutually entailed orienting may involve establishing orienting functions for particular stimuli, it may also involve establishing evoking functions. Consider, for example, a caregiver shouting loudly when a child approaches a dangerous stimulus (e.g., a plant with a tasty looking but poisonous berry). It is highly likely that that stimulus will acquire both strong orienting and (in this case aversive) evoking functions for the child (although this outcome could be explained in terms of Pavlovian fear conditioning, the term mutually entailed orienting/evoking is invoked here when it is directly involved in the acquisition of symbolic relations as defined within RFT). Furthermore, when an infant engages in mutually entailed orienting with a caregiver, even stimuli that are merely oriented toward by that caregiver (in the absence of any other cue or warning signal) may acquire relatively positive (approach) evoking functions for the infant. In addition, the ubiquitous nature of motivating variables likely impacts the degree to which orienting and evoking functions are established and are brought to bear in any given instance (e.g., establishing an aversive function for a tasty looking, but poisonous berry may be more difficult when a child is mildly food deprived). Thus, as with the entailment triangle introduced in the previous section, the inseparable nature of mutually entailed orienting, evoking, and motivating should be emphasized here. As the listening repertoire of the infant develops, the mutually entailed orienting and evoking functions of specific stimuli likely become related (in varying motivative contexts) to particular sounds (e.g., words) in an arbitrarily applicable manner. The term mutually entailed (orienting, evoking, and motivating) thus indicates that establishing a basic listener repertoire involves a child and caregiver engaging in cooperative acts while the caregiver frequently emits language-appropriate sounds.

At this point we should be clear that the concepts of mutually entailed orienting, evoking, and motivating may involve the cooperative behaviors sometimes referred to as mutual eye gaze, joint attention, and social referencing. Indeed, the latter three terms or concepts have sometimes been seen as important precursors to AARR itself. On balance, it seems important to distinguish between these non-RFT concepts and mutually entailed orienting, evoking, and motivating because recent RFT literature has begun to emphasize that human cooperation was the initial driver of the evolution of symbolic language and cognition (e.g., Hayes & Sanford, 2014; Wilson et al., 2014), a view that is consistent with some recent lines of thinking in evolutionary science itself (e.g., Wilson, 2007). To illustrate the argument we are making here, consider a dog who has been trained to fetch a particular object when their owner points or even gazes at the object and shouts "fetch". This interaction is clearly cooperative and involves some element of joint attention and perhaps social referencing. That is, in order for the dog to fetch the object, both the owner and the dog had to attend to the same object. Indeed, some research has also demonstrated that dogs can follow human pointing to find hidden food (e.g., Hare et al., 1998). For RFT, however, these interactions would not be considered mutually entailed orienting for the dog, unless they are shown to lead to the acquisition of symbolic relations by the dog, as defined within RFT (i.e., AARR, such as contextually controlled generalized symmetry responding). Without this outcome, the term mutually entailed orienting/evoking/motivating should not be applied to the dog's behavior in this example of joint attention (or social referencing).

Elaborating on the Concept of AARR in RFT

In proposing the concept of mutually entailed orienting and evoking (within a given motivational context), the very definition requires that as listening and speaking repertoires gradually develop, mutually entailed orienting and evoking functions for particular stimuli become related, in an arbitrarily applicable (symbolic) manner, to specific sounds (i. e., words) or signs. Initially, the symbolic stimulus relations that are established may be defined as mutually entailed relations, such as those involved between an object and a spoken word (e.g., when a child orients toward a toy teddy when it hears the word "teddy"). According to RFT mutually entailed relations provide the basis for the emergence of increasingly complex classes of stimulus relations, which have been divided into five levels of increasingly complex types of relating. Before considering this framework in greater detail, we will present a brief description of the generic RFT account for the establishment of different classes of AARR or relational operants, termed relational frames (i.e., basic units of combinatorially entailed relations), and their combination into increasingly complex networks of relations.

Basic Relational Framing

In establishing relational framing, evidence suggests that mutually entailed relational responding emerges first. That is, bidirectional derived relations between two stimuli first emerge (mutual entailment), followed by derived relations among three or more stimuli, referred to as combinatorial entailment (Lipkens et al., 1993). The concept of the basic relational frame typically involves both mutual and combinatorial entailment. Imagine the verbal community provides a child with direct reinforcement for looking at or pointing to the household pet dog upon hearing the dog's name (e.g., Snoop) or the word 'dog.' Direct reinforcement is also provided when the child emits appropriate responses such as saying "Snoop" or "dog" in response to appropriate contextual cues such as "what's the dog's name?", or for producing these responses simply when the dog is observed. Multiple exemplars of coordinating various stimuli with their names in various contexts thus serve to establish the relational frame of coordination such that direct reinforcement is no longer necessary for novel stimuli. In other words, derived coordination is established in the behavioral repertoire of the child. For example, imagine the child from the above example is subsequently presented with a picture of a hedgehog alongside the written word "hedgehog" and is told its name. Subsequently presenting the child with a picture of a hedgehog or the written word may then produce the response "That's a hedgehog!" from the child in the absence of direct reinforcement or prompting. That is, once the generalized frame of coordination is established (coordinating arbitrary sounds and pictures with objects in the environment), directly reinforcing a subset of relating behaviors should generate the complete set without the need for further training. Furthermore, when this pattern of relating (framing) has been established, the generalized response may be applied to any stimuli in the presence of appropriate contextual cues, such as the Crel "that is a".

While the above example focused on establishing derived coordination in the presence of appropriate contextual cues (e.g., the Crel "is a" to specify the relation between dog and 'Snoop'), other cues (e.g., "shorter than" and "longer than") would be similarly established across multiple exemplars to specify other patterns of responding (i.e., other relational frames). That is, the (arbitrary) relating that takes place is not based exclusively on the formal relations among and between stimuli but is also based on auxiliary contextual cues which determine appropriate relational responding. For example, consider an abstract sentence like "a hummingbird is smaller than a cat." Understanding that a hummingbird is smaller than a cat despite the word "cat" being audibly and physically shorter than the word "hummingbird" requires that the relationship between the stimuli be arbitrarily applicable rather than determined by physical characteristics like length. Across multiple relevant exemplars establishing appropriate patterns of AARR (i.e., relational frames), even propositions such as "X is smaller than Y" should produce appropriate derived relating (i.e., "Y is larger than X") despite no knowledge of what X and Y actually are.

Increasingly Complex Relational Networking

For RFT, more complex aspects of symbolic language and cognition (e.g., rule-following, etc.) are accounted for through the combination of relational frames into increasingly complex networks. For example, instructions or rules have been conceptualized within RFT as networks of relational frames typically involving coordination and other types of relations such as conditionality, under the control of appropriate Crel and Cfunc contextual cues, which transform the behavior control properties of stimuli specified in the network (Barnes-Holmes et al., 2001). Consider, for example, the instruction "if a dog is eating then do not pet it". This simple rule involves frames of coordination between the words "dog," "eating", and "pet" and an actual dog, the act of a dog eating, and a person petting a dog. The words "if" and "then" function here as contextual cues for a conditional relation between a dog eating and petting. Critically, the words "do" and "not" in the instruction transform the functions of a dog eating, such that they reduce the probability of engaging in any petting behavior while a dog is actually eating. In effect, the functions of a dog eating have been transformed by the network, such that it now reduces a specific behavior in this context. This type of conceptual analysis of rules as complex relational networks has also been modelled successfully in laboratory studies (e.g., O'Hora et al., 2004, 2014; see Harte, Barnes-Holmes, Barnes-Holmes, & Kissi, 2020, and Zapparoli et al., 2021, for recent overviews of this area).

Relating Relations

For RFT, more sophisticated levels of symbolic language and thought (e.g., analogical and metaphorical reasoning) are also readily explained by scaling up in the complexity of the derived relating involved. Consider, for example, a simple analogy "a chef is to food as a sculptor is to stone." Both 'chef' and 'food', and 'sculptor' and 'stone' are coordinated via the Crel "is to." In addition, both separate coordination relations are connected by another coordination relation via the Crel "as." Crucially, relating relations involves responding to one's own responding in that one relational response is related in some way to another relational response (e.g., the 'chef-food' relational response is coordinated with the 'sculptor-stone' relational response; see Stewart & Barnes-Holmes, 2004, for a review).

Relating Relational Networks

To scale up in the complexity of derived relating again, RFT proposes that entire relational networks are likely related to other entire relational networks. While the empirical research at this highly complex level of relating is currently very limited (but see Ruiz & Luciano, 2011), advanced verbal skills (e.g., comparisons of extended narratives and complex problem solving) seemingly involve this level of relating, but at the current time this suggestion remains almost entirely interpretive.

AARR and the Symbolic Self

A sizable body of research, both conceptual and empirical, has been conducted within RFT on a class of AARR referred to as deictic relational responding, a pattern of AARR considered critical for the emergence of the "symbolic self." For RFT, the gradual emergence of the symbolic self is seen as inextricably linked to what we described previously as mutually entailed orienting, evoking, and motivating because these are seen as providing the critical historical context for the emergence of a sense of self. That is, a history established by cooperation between two separate individuals (a caregiver and infant) in which the caregiver produces sounds (i.e., words) that subsequently participate in arbitrarily applicable relations with the infant (e.g., their name), the caregiver themselves (e.g., "Mommy"), and the stimulus they are orienting toward (e.g., a "toy"). A mother, for example, could pick up a child's toy, orient the infant toward the toy (i.e., hold the toy in front of the infant) and say "Look! *Mommy* has *your toy*. Do *you* want *Mommy* to give *you* the *toy*?" Of course, at first these words have no symbolic meaning for the child. However, the mutually entailed orienting (and associated evoking/motivating) observed here could be seen as a crucial part of the history that gradually establishes symbolic functions across numerous cooperative instances in the child's early life. In effect, the words involved in these kinds of statements and questions could be seen as acquiring appropriate symbolic functions and deictic relational responding thus emerges.

According to RFT, there are three key relations involved in deictic relational responding: the interpersonal (I-You) relation, the spatial (Here-There) relation, and the temporal (There-Then) relation. These three core deictic relations combine to form the most basic deictic relational frame. Responding in accordance with this frame thus involves locating oneself in space and time (e.g., I-Here-Now) relative to someone else (e.g., You-There-Then). For RFT, a child learning to respond in accordance with these relations is essentially seen as learning to relate 'self' to others in particular spaces at particular times. Imagine, for example, asking a very young child "What did you do this morning?" If, for example, the child responds with what a sibling is currently doing, they may be corrected and told "No, that's what your brother is doing here now. What did you do earlier?" Similar ongoing refinement of the three deictic relations thus shapes the child's responding so that they can respond appropriately to questions about their own behavior relative to others in the context of specific times and places (e.g., Barnes-Holmes, 2001; McHugh et al., 2004).

The concept of deictic relations highlights that a verbal human's sense of self is, in a sense, socially constructed. The construction of this self begins with the cooperative acts involved in mutually entailed orienting, evoking, and motivating. For example, when a caregiver orients an infant towards a toy, and asks the child "do *you* want the toy", the child may reach towards the object as part of that cooperative act, despite not yet understanding any of the sounds that the caregiver has produced. Nonetheless, the overarching evolutionary history of human cooperation involved in child-rearing has begun to establish the symbolic self from which all events will be viewed. When considered in this way, the symbolic self is another term for having a perspective, as experienced by a verbal human. In a technical RFT sense, therefore, it is not possible for a verbal human to experience or know the world nonsymbolically (at least in a pure sense of 'nonsymbolic'), once a history of mutually entailed orienting/evoking/motivating has begun to establish arbitrary relations with I-You, Here-There, and Now-Then.

Deictic relating is seen as relatively advanced in that it involves responding to one's own responding. It is highly likely, therefore, that it is involved in relating relations and relating complex relational networks to other complex relational networks. To put it another way, a person would find it difficult to relate two relational responses to one another if they could not "locate" those relational responses in a specific time and space. Indeed, Kavanagh et al. (2020) recently elaborated this argument in their conceptual RFT analysis of a false belief perspective-taking task.

Before moving on we should emphasize that in providing "everyday" examples of scaling up in levels of complexity (e.g., rules, analogical reasoning, the verbal self, etc.), those examples are *not* explanations for AARR, but rather AARR is the technical (behavior-analytic) explanation for the examples. For instance, RFT may define a rule as a relational network, but not all networks are rules. To put it another way, the concept of the relational network is the technical concept used to explain how a rule (a nontechnical concept) controls the behavior of a listener, but it would be a mistake to conclude that the everyday concept of a rule has any explanatory weight inside RFT as a scientific account itself (Harte & Barnes-Holmes, 2021a). The functional-analytic abstractive nature of RFT, as an explanatory approach to explaining human language and cognition, is perhaps best understood or appreciated through the lens of the new framework, mentioned earlier, to which we now turn.

Conceptualizing and Analyzing the Dynamics of AARR: A Hyper-Dimensional, Multilevel (HDML) Framework

According to RFT, the establishment of AARR allows for the evolution of increasingly complex forms of relational responding (discussed above), 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). The development of increasingly complex types of AARR has recently been formalized within some of the RFT literature as a hyper-dimensional, multilevel (HDML) framework. At the present time, the HDML specifies five levels of relational development: (i) mutually entailing, (ii) combinatorial entailing, (iii) relational networking, (iv) relating relations, and (v) relating relational networks. The framework also emphasizes the dynamic nature of the relating activity that may occur along four dimensions. The four dimensions are coherence, complexity, derivation, and flexibility, which are briefly defined below (for more detailed treatments see Barnes-Holmes et al., 2017, 2018, 2020).

Coherence refers to the extent to which a pattern of AARR is consistent (coheres) with previously established patterns. For example, imagine you are told that a snake is longer than a centipede, and subsequently told that a centipede is shorter than a snake. It is likely that the second statement would be considered coherent or consistent with the first. Coherence here would be high given that the overall pattern ($X > Y = Y < X$) is consistent with the way in which these verbal relations have been established by the wider verbal community. In other words, there are few instances in which an English-speaking listener would reinforce, or not correct, the statement, "if X is bigger than Y, then Y is bigger than X").

Complexity refers to the density or detail involved in a particular pattern of AARR, such as types and/or number of relations involved. For example, a mutually entailed relation of coordination may be seen as less complex than a mutually entailed temporal relation because the former involves only one type of relation (e.g., if X is the *same as* Y, then Y is the *same as* X), whereas the latter involves two types (if X occurs *before* Y, then Y occurs *after* X).

Derivation refers to the extent to which a particular pattern of AARR has been emitted previously. Each time a pattern of AARR is derived, its derivation reduces as it acquires its own history beyond the original derivation. Imagine, for instance, that a person learns that snakes are longer than centipedes, and thus derives that centipedes are shorter than snakes. The first time that the 'centipedes-shorter-than-snakes' relation is derived, it is derived "directly" from the 'snakes-longer-than-centipedes' relation. However, as the person continues to relate centipedes as shorter than snakes, that relational response acquires its own history, rendering it less and less derived from the original relation (regardless of whether it is reinforced directly). As an aside, Barnes-Holmes et al. (2016) noted that

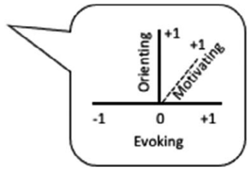
> The concept of derivation may be used in two ways within RFT. First, it may indicate that arbitrarily applicable relational responding is derived from a history of operant learning, or multiple-exemplar training. That is, the ability to engage in such relational responding is derived from a continuously growing source of prior learning. When derivation is used in this first sense, it may be seen as increasing across repeated instances because the source of the derivation (i.e., the generalized operant class itself) grows with every reinforced exemplar. The second way in which derivation may be employed within RFT is in the sense that derived relational responding involves "arriving at a conclusion" based on a relation, subset of relations, or multiple exemplars of such relations. More informally, it involves, "reasoning," "inferring," or "deducing" from a limited source of prior learning. Critically, it is only in the second sense that derivation may be interpreted as decreasing across multiple instances of a particular pattern of arbitrarily applicable relational responding. (p. 126)

Flexibility refers to the extent to which a given pattern of AARR may be modified by current contextual variables. As a simple example, imagine a young child who is asked to respond with the wrong answer to the question, "Which is longer, a snake or a centipede?" The faster the child responds with "centipede", the more flexible the AARR (see O'Toole & Barnes-Holmes, 2009). Critically, flexibility is always seen as context dependent—thus if the child had been told previously *not* to give a wrong answer when asked to do so, it would be difficult to use a

Table 1

An Illustration of the HDML Framework

Levels	Dimensions			
	Coherence	Complexity	Derivation	Flexibility
Mutual Entailing	*Analytic Unit 1*	*Analytic Unit 2*		
Relational Framing				
Relational Networking				
Relating Relations				
Relating Relational Networks			*Analytic Unit 19*	*Analytic Unit 20*



Note. The five levels and four dimensions of arbitrarily applicable relational responding, appreciating the critical role of orienting and evoking functions, and motivating variables. Motivating is represented by a broken line because its impact is inferred based on changes in orienting and evoking functions. Each level intersects with each dimension giving rise to 20 functional analytic-abstractive units of analysis. The table aims to capture the dynamic and hyperdimensional nature of AARR as relating, orienting, evoking and motivating (ROE-M).

correct or wrong answer as an indication or measure of flexibility.

The intersections between the levels of relational development and the four dimensions yield 20 units of analysis, thus producing the HDML framework (Table 1). Table 1 also illustrates the ROE-M which emerges from the framework. The ROE-M provides a way of conceptualizing human psychological events as they unfold in real time (see also Harte & Barnes-Holmes, 2021b). Specifically, each cell of the HDML contains an inverted ‘T’ with a third dashed line. This symbol represents orienting and evoking functions, and the impact of motivating variables, which may occur within each of the 20 units of AARR. Orienting refers to the extent to which a stimulating event is noticed or “stands out” in the wider context. Conceptually, orienting within the HDML is seen as lying on a continuum, on the vertical axis, from complete absence (0) to strongest orienting response possible (1). Evoking refers to the extent to which a stimulating event is deemed to be appetitive versus aversive. Evoking is also seen as lying on a continuum, on the horizontal axis, from the strongest aversive response possible (-1) to the strongest appetitive response possible (+1) with 0 representing the absence of either an aversive or appetitive reaction. Motivating is represented with the broken line, which is scaled from 0 to 1, indicating the putative strength of motivational variables, which interact with orienting and/or evoking functions, and indeed relating, in a dynamical manner (see below).

The core or fundamental unit of analysis for AARR is thus defined as *relating*, *orienting*, and *evoking* within a given *motivational* context (i.e., the ROE-M). In conceptualizing human verbal or symbolic responding in terms of the ROE-M, it is critical to appreciate that the four dimensions of the HDML provide the historical and current contextual variables (antecedent and consequential variables) that allow for the prediction-and-influence of the ROE-M itself. For example, a “reinforcer” does not simply strengthen a pattern of “ROE-Ming” but is seen as perturbating the dynamic interplay among the elements within the ROE-M unit. Thus, awarding a point on a computer screen during an operant task for a verbal human does not simply increase key-pressing, but may impact upon previously emitted relational networks (e.g., self-generated rules) that then impact upon the extent to which particular stimuli control orienting and evoking functions in the experimental task. Indeed, the delivery of an "unexpected" point, could serve to undermine the reinforcing value (an evoking function) of points per se (e.g., see Harte, Barnes-Holmes, Moreira et al., 2021, p.477). As an aside, it is worth noting that "rethinking" the concept of reinforcement itself in behavior analysis is not unique to RFT. For example, Baum (e.g., 2018) has argued recently that reinforcement does not simply strengthen operant behavior, but selects and directs it, and Cowie (e.g., 2018) has argued that a history of reinforcement exerts both prospective *and* retrospective control over behavior.

The RFT analysis outlined above proposes that most if not all human psychological events (for verbal humans) involve the ROE-M. As an illustrative example, a mutually entailed relation (e.g., "Vaccines are wonderful") may be conceptualized as varying in coherence, complexity, derivation, and flexibility. In general terms, the relation between vaccines and wonderful may be relatively high in coherence if the statement coheres with similar assertions (e.g., "the Covid-19 vaccine will save millions of lives"); relatively low in complexity if understanding the statement involves a limited number of other relational responses (e.g., the word "vaccine" was applied directly to a positive event, such as the end of Covid-19 lockdown); relatively low in derivation (e.g., if similar statements have been heard many times in the past, such as "vaccines have eradicated many deadly diseases"); and low in flexibility (e.g., if it is difficult to modify or "challenge" the perceived truth of the statement, such as readily dismissing "anti-vax" arguments). Critically, this relational activity is seen to interact in a nonlinear and dynamical manner with the orienting and evoking functions of stimulating events for humans as they navigate their environments. For example, a passing comment ("I can't wait to get my Covid-19 vaccination") may increase orienting and (positive) evoking functions for an expected text notification for an appointment to receive a Covid-19 vaccination. Various motivational variables will likely impact on the relative strengths of the orienting and evoking functions, such as personal health risk factors associated with contracting Covid-19 (e.g., age, weight, pre-existing medical conditions, etc.).

It is important to emphasize the inseparable, interactive and nonlinear nature of ROE-Ming that the HDML aims to capture. The example provided above of a positive comment about vaccines (i.e., they are wonderful) could increase orienting and evoking functions for news reports about the roll-out of the Covid-19 vaccine in your area. In contrast, imagine that no such comment was made. You may be less likely to orient as strongly and as positively towards such news reports, but may still engage in some level of relating, such as responding in accordance with a relational network that functions as a self-rule (e.g., "I must keep an eye open for any information of the roll-out of the Covid-19 vaccine in my area"). And of course, the nature and relative strength of the relating and orienting/evoking will be moderated by motivational variables. In essence, the concept of the ROE-M is designed to capture the constant, dynamical, and nonlinear nature of the core unit of responding that characterizes human psychological events. From an RFT perspective, the set of relational abilities, and associated orienting and evoking functions (interacting with motivational variables) contained within the ROE-M, as a unit of analysis, appear to be core defining characteristics of the human species, which allow us to navigate and react within our physical and social environments in increasingly sophisticated and powerful ways.

As noted earlier in the current article, it seemed useful to first lay out the *conceptual* developments that have emerged in some of the RFT literature in recent years without constantly referring to, and explaining, the empirical research that drove the bulk of those conceptual advances. Having provided the narrative on the conceptual developments we will now focus more on the empirical work that generated much of the conceptual development we have described above. It is important to emphasize that the empirical developments summarized below served to *generate* the HDML and the ROE-M, but in many cases could not be considered formal tests of these new conceptual tools. Indeed, in the case of the HDML framework it should not be considered a *model* that will yield to a formal hypothetico-deductive type of analysis. Rather, it is offered as a framework that may (or may not) help to guide the scientific behaviors of RFT researchers, and perhaps others, going forward. Specific models may emerge from the framework that could be tested, but the framework itself should be seen as largely an inductive scientific device.

Empirical Developments that Led to the HDML and the ROE-M

Shortly after publication of the seminal volume on RFT (Hayes et al., 2001) a colleague asked whether it was possible to capture relational framing "in flight." This was interpreted by the first author of the current article as asking how we might measure the relative strength or probability of specific patterns of AARR. The initial response was the development of what came to be known as the Implicit Relational Assessment Procedure (IRAP).

In developing the IRAP, two separate methodologies were combined. The first was an RFT-based procedure developed to train and test multiple stimulus relations called the Relational Evaluation Procedure. The second was the Implicit Association Test (IAT), which had been developed by social-cognition researchers for measuring what they conceptualize as associative strengths in memory (Greenwald et al., 1998). Combining the two procedures appeared to offer a measure of the strength of natural verbal relations, or AARR (Barnes-Holmes et al., 2008). However, as a result of its close connection to the IAT, research with the IRAP quickly became dominated by a focus on so-called implicit attitudes and implicit cognition more generally. Although this strategy provided a useful means of assessing the validity of the IRAP in a number of contexts (Vahey et al., 2015), it also somewhat detracted from the original purpose of its conception: studying the dynamics of AARR. The critical point here is that the IRAP was initially developed within RFT as a method for assessing the relative strength of specific patterns of AARR because they are seen as providing the functional-analytic units of symbolic language and cognition. In what follows, we will concentrate on IRAP research that remained focused on its initial purpose rather than as a "mainstream" instrument for measuring so-called implicit cognition.

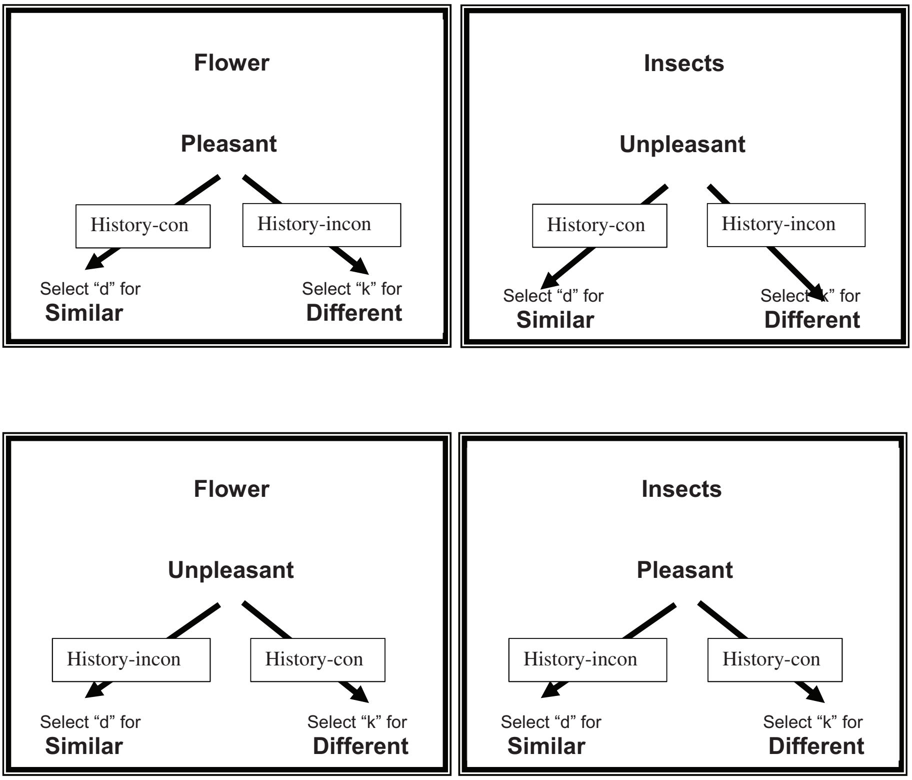
The IRAP: Procedural and Analytic Overview

The IRAP is a computer-based task in which an individual responds to a series of screens containing combinations of verbal stimuli (i.e., stimuli that have acquired their functions based, at least in part, on a history of AARR). Label stimuli—such as flower and insect—appear at the top of the screen (see Fig. 2). Target stimuli—such as pleasant, good, unpleasant, and bad—appear in the middle of the screen. On each trial, two response options are presented that specify particular relationships between label and target stimuli. For example, 'flower' and 'pleasant' might appear on a given trial with the response options 'similar' and 'different.' In this case, participants would be required to relate flowers as similar to, or different from, pleasant. The IRAP operates by requiring opposite patterns of responding across successive blocks. For example, 'flower' and 'pleasant' would require the response 'similar' on one block and 'different' on the next. This was based on the assumption that, all things being equal, the more frequently reinforced (and thus more probable) response pattern, or one that was relationally coherent with it, would be emitted more readily (Barnes-Holmes et al., 2010). To increase the likelihood that the more probable response is observed, responding on the IRAP is typically placed under time pressure (e.g., participants are required to respond within 2000 ms on each trial).

Within the verbal community, certain relational responses are more likely to be reinforced than punished (e.g., affirming that flowers are pleasant), whereas others are more likely to be punished than reinforced (e.g., denying that flowers are pleasant). Thus, the more readily emitted pattern of responding should be indicative of the natural contingencies operating in the wider verbal community (i.e., using stimuli that are already "meaningful" to the participants). Broadly speaking, the IRAP is scored by subtracting the mean response latency for one pattern of responding from the mean response latency of the opposite pattern of responding. The resulting difference, if any, is deemed to be reflective of the differential reinforcement for the two patterns of responding (or relationally coherent patterns) in the preexperimental history of the individual. In most IRAP studies, four difference scores are calculated, one for each of the four trial types typically presented within the IRAP (e.g., *flowers-pleasant*; *flowers-unpleasant*; *insect-pleasant*; *insect-unpleasant*). The "predicted" pattern of responses might thus be faster

Figure 2

An Example of the Label-Target Combinations Presented On-Screen Comprising Each of Four Trial-Types in an IRAP



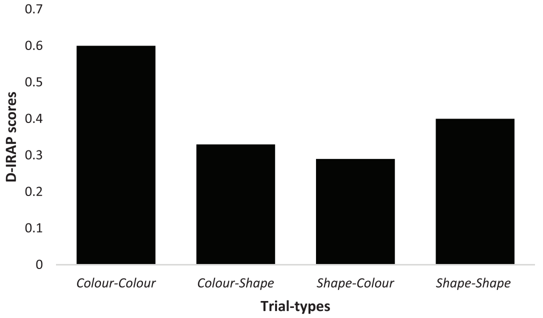
Note. The arrows and boxes indicating correct responding on history-consistent and -inconsistent trials would not appear on the screen.

responses when confirming, rather than denying, that flowers are pleasant and insects are unpleasant; and denying, rather than confirming, that flowers are unpleasant and insects are pleasant.

One strategy to test the basic assumption underlying the IRAP might involve using stimuli with virtually no strong emotional functions. For example, if a study employed the words "shapes" and "colors" as labels with relevant examples (square, circle, red, blue, etc.) as targets, then surely the four trial-type scores (shape-shape; shape-color; color-shape; color–color) would be roughly equal. However, even this simple assumption proved to be faulty (e.g., Finn et al. 2016; O'Shea et al., 2016). For example, Finn et al. (2018) employed an IRAP of this very nature (in this study the response options were 'True' and 'False' rather than 'similar' and 'different'; see below for additional detail). Results showed that even though the effects were in the predicted direction (shorter latencies during history-consistent blocks relative to history-inconsistent blocks), the effect for the *color–color* trial-type was significantly larger than the other three (*color-shape*, *shape-color*, and *shape-shape*; see Fig. 3). The authors suggested that the smaller effect sizes for the *color-shape* and *shape-color* trial-types could be readily explained by the fact that history-consistent blocks of trials required selecting 'False' rather than 'True'. That is, if there was an inherent response bias toward confirming relations (rather than disconfirming them), then on trials with False as the correct response, reduced effect sizes would be expected (i.e., for a noncoherent relation, such as *shape-color* or *color-shape*). This explanation could not, however, account for the larger effect observed for the *color–color* trial-type relative to *shape-shape* because both trial-types required participants to respond 'True' during history-consistent blocks. In attempting to explain why this disparity

Figure 3

Broad Pattern of Trial-Type Effects Produced by the 'Shapes And Colors' IRAP



Note. While the Finn et al. (2018) study divided participants into two experimental groups based on IRAP experience, this is not represented in the current figure. Rather the general pattern of effects is illustrated.

emerged between two trial-types, which in principle should produce near identical effects, the authors argued that the color words employed appeared with higher frequency in natural language than the shape words (Keuleers et al., 2010). Thus, the color words produced a stronger orienting response (e.g., were more "attention-grabbing") than their shape counterparts for the average participant because color words occur in everyday discourse far more often than shape words. Of course, there may be other reasons that the concept of colors (and relevant exemplars) may be more salient than the shape stimuli, but the term "orienting functions" was employed to recognize that even relatively bland stimuli may differ in their orienting properties (i.e., in the extent to which they grab attention or 'stand out', etc.).

The Potential Role of Orienting Functions and the Beginnings of a Model for Analyzing Differential IRAP Effects

In grappling with the differential trial-type effects described above and developing an appropriate formal explanation, a differential arbitrarily applicable relational responding effects (DAARRE; pronounced "dare") model was proposed (Finn et al., 2018). According to the model, such differential trial-type effects may be explained by the extent to which the *Cfunc* and *Crel* properties of the stimuli employed within a given IRAP cohere (overlap) with certain properties of the response options across blocks of trials. In addition to these two technical terms, *Crel* and *Cfunc*, the concept of the relational coherence indicator (RCI) has been proposed within the RFT literature. The need to add this concept became apparent in attempting to define the functions of specific response options employed within the IRAP (see Maloney & Barnes-Holmes, 2016, for a more detailed discussion of RCIs). For example, many IRAP studies employ the RCIs, "True" and "False" as response options—in contrast to *Crels*, such as "similar" and "different," which were used in early IRAP research. The RCIs, "True" and "False" were employed to specify relational coherence or incoherence, in general, between label and target stimuli, rather than specific relations, such as 'similar' and 'different.' An illustration of the basic DAARRE model applied to the 'shapes and colors' IRAP is presented in Figure 4.

This DAARRE model analysis highlights three key sources of behavioral influence: (1) the relation between the label and target stimuli (the *Crel* property); (2) the orienting property of the label and target stimuli (a *Cfunc* property); and (3) the coherence/incoherence property of the RCIs, in this case "True" and "False".[3] As mentioned previously, the two critical trial-types in this analysis were the *color–color* and *shape-shape* trial-types. As can be seen from the figure, the Cfunc property for 'colors' is denoted as positive ('+' sign) while the Cfunc property for 'shapes' is denoted as negative ('-' sign). This is to reflect the suggestion made above that the color stimuli likely possess stronger orienting functions

[3]As noted above, the terms *Crel* and *Cfunc* are used in RFT to refer to two types or classes of contextual control. When the term Crel or Cfunc *property* is employed it simply refers to the specific behavioral function that a stimulus has acquired via the history of these two types of control. For example, the coordinating relation between the word apple and actual apples is its Crel property (established at some point under the control of an appropriate *Crel*); any other behavioral function actualized by the word apple, such as the extent to which it "grabs your attention" or evokes a positive or negative reaction in an IRAP, is defined as its Cfunc property. The contextual control in this case may be relatively complex in that it depends on a history of interacting with actual apples and an appropriate history of AARR that allows specific properties of apples to be selected. Following on from the example of Cfunc control in the main text, "imagine the taste of an apple", an appropriate response can only occur if the listener has previously eaten an apple and learned to respond appropriately to words such as "imagine" and "taste."

relative to the shape stimuli given their differential frequency in natural language. It is important to note, however, that the negative labelling for shapes should not be interpreted as specifying negative orienting functions, but rather orienting functions that are weaker relative to that of colors. Returning to Figure 4, the relations between label and target stimuli are also denoted by plus or minus signs, in this case to indicate the extent to which they do or do not cohere based on the participants' relevant learning history. As such, the color–color and shape-shape relations are denoted by plus signs to indicate coherence, while the color-shape and shape-color relations are denoted by minus signs to indicate incoherence. Finally, both response options are also denoted by plus and minus signs to specify their functions as either coherence or incoherence indicators. That is, in natural language, 'True' would typically be employed to indicate coherence (+) and 'False' to indicate incoherence (-).

To appreciate the explanation offered by the DAARRE model for the differential trial-type effects found in Finn et al. (2018), first consider the *color–color* trial-type and observe that all Cfunc and Crel properties are labelled with '+' signs. In addition, the RCI deemed correct on history-consistent trials (i.e., 'True') is also labelled with a plus sign, making this trial-type the only one with four plus signs. Thus, this trial-type may be considered maximally coherent during history-consistent blocks of trials. During history-inconsistent blocks of trials, however, the RCI that is deemed correct (i.e., 'False') is labelled with a minus sign meaning there is no coherence between the properties of the Crel and Cfuncs (all plus signs) with the required RCI (minus sign). According to the DAARRE model analysis, it is this clear contrast in levels of coherence across blocks of trials that results in a relatively large IRAP effect for this trial-type.

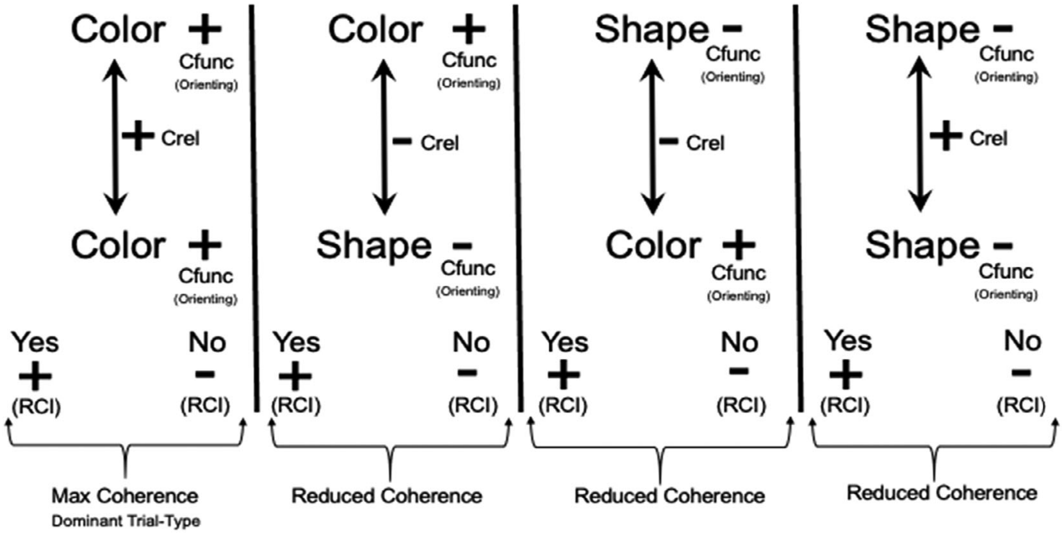
Next, consider the *shape-shape* trial-type. During history-consistent blocks of trials, although participants are required to choose the same RCI as for the *color–color* trial-type ('True'), the property of this RCI (a plus sign) only coheres with the Crel property between the label and target stimuli but not with the Cfunc properties of both stimuli (both minus signs). During history-inconsistent blocks, however, the Cfunc properties of the label and target stimuli and the required RCI ('False') do cohere (three minus signs), albeit no longer with the Crel property. Thus, the disparity in coherence across these two trial-types (*color–color* and *shape-shape*) between history-consistent and history-inconsistent trials is unequal.

Finally, upon inspection of the remaining *color-shape* and *shape-color* trial-types, the difference in coherence across history-consistent and -inconsistent blocks of trials is further reduced when compared to the *color–color* trial-type (two relative to four plus signs). This, at least in part, could explain the dominance of the *color–color* trial-type over the remaining three. A number of recent studies have provided experimental evidence in support of this DAARRE model explanation and the role of orienting functions (Bortoloti et al., 2020; Finn et al., 2018, Experiment 3; Finn et al., 2019; Pinto et al., 2020; Schmidt et al., 2021). We will return to the important issue of empirical testing in greater detail later in the article.

At this point, it seems important to explain that the use of the term "coherent" within the DAARRE model is consistent with the definition provided earlier in the context of the HDML (i.e., "the extent to which a pattern of AARR [which involves both Crel and Cfunc properties] is consistent with previously established patterns"). The term does *not*, therefore, only apply to the Crel properties of stimuli (e.g., X > Y coheres with Y < X), but also to the Cfunc properties of stimuli, which include RCIs. In the case of the 'shapes and colors' IRAP described above, the *color–color* trial-type is thought of as maximally coherent because relatively strong confirmatory responses are involved for all of the critical responses during a history-consistent block of trials. In making this suggestion, we are assuming that most participants (if not all) would be susceptible to a general confirmation bias effect (e.g., Nickerson, 1998). From an RFT perspective, such a bias is generated by a learning history involving a higher frequency of "confirming responses" for stimuli and events that are functionally similar rather than for those that are functionally dissimilar. In principle, it would be possible to test this suggestion and manipulate coherence within the IRAP. For example, by exposing participants to a 'shapes and colors' IRAP that presented

Figure 4

The DAARRE Model as it Applies to the 'Shapes And Colors' IRAP



Note. The '+' and '-' symbols denote the relative positivity of the Cfunc property of each label and target stimulus, the relative positivity of the Crel property, and the relative positivity of the relational coherence indicator (RCI) in the context of the other Cfuncs, Crels and RCIs. Figure previously published in *The Psychological Record*. Reproduced with permission of Springer.

the *shape-shape* trial-type far more frequently than the other three trial-types (across numerous sessions), the pattern of Crel and Cfunc properties for the stimuli involved in this trial-type would likely increase in coherence (thus over-riding the standard confirmation bias effect). That is, there would be greater functional overlap between the response pattern on the *shape-shape* trial-type and the dominant pattern observed during previous sessions for that IRAP. In our view, this all-embracing functional definition of coherence is necessary when the ROE-M (which involves both Crel *and* Cfunc properties) is defined as the generic unit of analysis (as co-determined by motivational variables).

Extending the DAARRE Model with the Potential Role of Evoking Functions

Appreciating the potential importance of evoking functions in AARR became apparent to the Odysseus research group in IRAP studies reported by Leech et al. (2016, 2017). The IRAPs employed in both studies comprised 'cute' puppies or kittens versus aggressive-looking spiders as label stimuli, approach (e.g., "I can pick it up") versus avoidance ("I need to get away") phrases as target stimuli, and the RCIs "True" and "False" as response options. The patterns of effects produced on the IRAPs in both studies were broadly intuitive in that positive response biases were produced on the two pet trial-types (*pet-approach-true* and *pet-avoid-false*) and a negative response bias was produced on one of the spider trial-types (*spider-avoid-true*). Interestingly, however, the pattern produced on the *spider-approach* trial-type was in a counter-intuitive direction. That is, one might predict a negative response bias on this trial-type on the assumption that participants would not readily approach spiders in the natural environment. Counter-intuitively, however, a positive response bias was produced (i.e., participants tended to respond "True" more quickly than "False" to spider images and approach descriptors). Curiously, this trial-type significantly correlated with participant performances on a behavioral approach task that involved approaching a live spider. Specifically, the more readily participants chose "True" than "False" on the *spider-approach* trial-type, the more likely they were to approach a live spider. Thus, although the direction of the response pattern on this trial-type appeared counter-intuitive, it predicted *actual* behavior.

In grappling conceptually with this finding, Barnes-Holmes et al. (2020) suggested that it is possible two Cfunc properties were involved in determining participant responses—orienting *and* evoking. To appreciate this interpretation, first consider the label stimuli employed within the IRAP. The "cute" pet pictures would likely possess strong (appetitive) evoking functions but somewhat weaker orienting functions given their inherent lack of threat. On the other hand, relative to the pet pictures, the spider pictures likely possessed strong (aversive) evoking functions *and* strong orienting functions given that they could be seen as potentially threatening or dangerous. Now consider the target stimuli; while the approach and avoidance phrases may not have differed in the relative strength of their orienting functions, their evoking functions probably differed such that the avoidance descriptors likely possessed relatively aversive functions and the approach descriptors relatively appetitive functions.

Finally, the particular functions of the stimuli that dominated for any given participant may have differed relative to their self-reported spider fear. When interpreted through the lens of the DAARRE model, therefore, it is likely that for participants relatively low in self-reported spider fear (i.e., participants for whom spiders were not particularly appetitive or aversive) the orienting functions of the spider pictures dominated over the evoking functions. By comparison, it is likely that for participants relatively high in self-reported spider fear (i.e., for whom spiders were seen as mildly threatening), the aversive evoking functions of the spider stimuli dominated over the orienting functions. When viewed in this way, responding with "True" more quickly than "False" would be a relatively coherent response for participants considered low-fear but less coherent for those considered high-fear. An illustration of the DAARRE model applied to the *spider-approach* trial-type is presented in Figure 5.

As shown in Figure 5, the Crel between spiders and approach is denoted as negative (i.e., noncoherent) because most people would not report eagerly approaching spiders. "False", therefore, would be considered the correct response on history-consistent trials. However, as explained above, a relatively strong orienting function is likely established for spiders for the low-fear participants but a relatively strong evoking function for high-fear participants. For low-fear participants, therefore, the dominating Cfunc property for spiders (orienting) is positive, as is the Cfunc property for the approach target stimulus (evoking), and both cohere with the positive ('True') RCI. In the case of high-fear participants, however, the dominating Cfunc property for spiders (evoking) is negative but positive for the approach target stimulus (evoking). As such, for high-fear participants one of the Cfunc properties coheres with the 'True' RCI (positive) and the other coheres with the 'False' RCI (negative). While admittedly speculative and post-hoc, if correct this interpretation would explain why performance on this particular trial-type reliably predicted actual approach behavior in both Leech et al. studies despite its seemingly counter-intuitive direction (2016, 2017).

Before continuing it seems important to mention that the question still remains why the effect produced on the *Spider-Approach* trial-type tended to be opposite in direction than expected (i.e., choosing "True" more quickly than "False")? Indeed, this type of effect has been reported in other studies using different stimuli (e.g., Kavanagh et al., 2019). That is, when comparing the performance of both two trial-types that require choosing "False" during history-consistent blocks of trials, the IRAP effect produced for the *negative–positive* trial-type is often weaker than that produced for the *positive–negative* trial-type. Given that both of these trial-types require the same response option (RCI), how might this difference be explained? Appeal to the DAARRE model may be useful here. The reader is referred back to Figure 5 in which the Cfunc properties of the target stimulus and "True" RCI on the *Spider-Approach* trial-type cohere with each other (both plus signs). In contrast, on the *Pet-Avoid* trial-type the Cfunc properties of the target stimulus and "False" RCI cohere. Assuming that the spatial contiguity between the target and response option helps determine participants' responses, the observed difference in trial-type effects seems to make more sense. In other words, assuming that participants generally read each trial from top to bottom, they may react to the *Pet-Avoid* trial-type as "*Yes-No-No*", but to the *Spider-Approach* trial-type as "*No-Yes-No*". If, therefore, participants find it easier to choose an RCI that is functionally similar (rather than dissimilar) with the target stimulus they have just seen, then the weaker (or "counter-intuitive") effect for the *Spider-Approach* trial-type is readily predicted (see Finn et al., 2019; Kavanagh et al., 2019).

The Impact of Motivating Variables

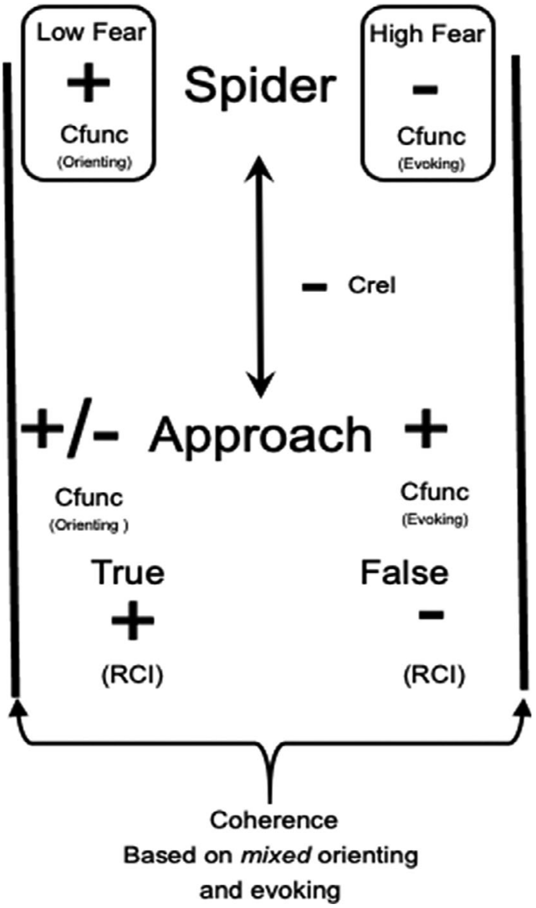
As noted earlier in the article, the ROE-M highlights the ubiquitous nature and influence of motivating variables in determining the orienting and evoking functions of stimuli. The importance of motivating variables was already recognized, in the seminal RFT volume with the concept of augmenting, and indeed within behavior analysis more generally (e.g., Michael, 1993, 2007; Skinner, 1953, 1957). However, the concept of augmenting was typically applied to rule-governed behavior within RFT and its relative precision, as a technical term, has been questioned (see Harte & Barnes-Holmes, 2021a, for an extended discussion). On balance, the concept of motivation, per se, remains important, as reflected in its inclusion inside the ROE-M. Indeed, a recent study explored the impact of three different motivative conditions on IRAP performance (Gomes et al., 2020). Researchers here used pepper sauce to increase the appetitive functions of water-related stimuli presented within an IRAP. Results showed that when participants ingested two drops of pepper sauce before completing the IRAP, the effect for the water-positive trial-type was substantially and dramatically larger than the effect for this trial-type for participants who only ingested a single drop of pepper sauce or did not ingest any. It thus appeared that the (appetitive) evoking functions of the water-related stimuli increased differentially when a motivating variable for water was introduced and manipulated. Given this finding, and the importance of motivating variables already evident from the wider behavior analytic literature, it seems wise to assume that motivating variables are a ubiquitous part of the field of symbolic interactants, ever present in determining the properties of any given instance of ROE-Ming.

Emerging Empirical Research Generated by the HDML

What we have presented thus far has aimed to provide a succinct overview of conceptual and empirical work generated by the Odysseus research group and its collaborators. In doing so, we have introduced new terms and

Figure 5

The DAARRE Model as it Applies to the Spider-Approach Trial-Type in Leech et al. (2016, 2017)



Note. "Low Fear" and "High Fear" indicate the Cfuncs that likely dominate for participants who were low (orienting) versus high (evoking) in spider fear. The "+/-" symbol is used to denote the assumption that the orienting functions of "approach" relative to "avoidance" target stimuli would likely not differ dramatically in this particular IRAP. Figure previously published in *The Psychological Record. Reproduced with permission of Springer.*

concepts that extend beyond those presented in the 2001 seminal RFT volume. We fully recognize that some of the material could be considered somewhat (or in certain cases even wildly) speculative. Nonetheless, it is important to emphasize that the new terms and concepts were generated by effects from the experimental laboratory that could not be explained readily by existing RFT concepts. More importantly, our objective was to generate novel ideas for future research that may contribute to RFT and behavior analyses of human language and cognition more generally. In this regard, certain studies could be seen as testing some of the new concepts contained within the HDML, rather than being responsible for generating them. For example, Finn et al. (2019) attempted to "engineer" the differential trial-type effect found in the Finn et al. (2018) study described above, and use the DAARRE model to predict the pattern of effects that would be found. Specifically, a pretraining phase was employed to establish confirmatory ("True") functions for one of two relatively innocuous stimuli (pictures of forks and spoons). Participants were divided into two groups, one in which the pretraining involved establishing confirmatory functions for the fork stimulus and the other in which confirmatory functions were established for the spoon stimulus. Consistent with the DAARRE model, stronger orienting functions would likely be established for the stimulus involved in the confirmatory pretraining (i.e., forks in one group and spoons in the other) thus producing a stronger IRAP effect for the trial-type that employed the pretrained confirmatory stimulus. Results were in accordance with this prediction. Furthermore, the authors reported that the effect "…was shown by 28 of the 40 participants, a result that would be observed, by chance alone, in less than 1 in 100 instances" (p. 434).

Recent research by Pinto et al. (2020) employed eye tracking technology to assess the role of orienting functions for smart phone application icons (WhatsApp vs. Messenger) and emotional faces (happy vs. angry). The study also employed an IRAP to assess response biases for these stimuli. Critically, the researchers assessed whether the response biases produced on the IRAP corresponded with the differential orienting responses on a visual search task, which monitored eye movement. The pattern of IRAP effects that emerged showed that the strongest trial-type effect was produced when the WhatsApp icon and happy face were presented as label and target stimuli. Eye-tracking data from the visual search performance task revealed that participants were faster to find (i.e., orient towards) the WhatsApp icon than the Messenger icon, and the happy faces rather than the angry faces. This finding was entirely consistent with the DAARRE model prediction that the differential trial-type effect may be driven, in part, by stronger orienting functions for both the label and target stimuli presented within the IRAP. Indeed, the authors concluded that the current data provide support for the potential importance of stimulus orienting functions and the DAARRE model more generally.

In another recently published article, Bortoloti et al. (2020) reanalyzed previously published IRAP data (Bortoloti et al., 2019) in light of work emerging from the DAARRE model. In the initial 2019 study, participants were first provided with a task that aimed to establish two equivalence classes, each involving an emotional face (a happy face in one class and a negative face in the other) and pseudowords. These stimuli were subsequently inserted as label and target stimuli into an IRAP. Despite each equivalence class being trained to the same performance criteria, differential trial-type effects emerged. Consistent with the DAARRE model, the trial-type in which the Cfunc and Crel properties of the stimuli and response options overlapped the most (i.e., responding 'True' more quickly than 'False' to a happy face and "happy" pseudoword) produced the largest IRAP effect. Interestingly, the same type of counterintuitive effect reported by Leech et al. (2016, 2017) and explained above in terms of the DAARRE model (see Kavanagh et al., 2019) was highlighted in the reanalysis of Bortoloti et al.'s data. Specifically, participants responded 'True' more quickly than 'False' when presented with a negative pseudoword and happy face. The authors hypothesized that the functional overlap between the functions of the happy face and the response option 'True' may have increased the likelihood that participants would respond in this direction. Indeed, the authors acknowledged that their explanation was consistent with the DAARRE model explanation for the same type of effect, as offered by Kavanagh et al. (2019). Furthermore, in a very recent follow up study, Schmidt et al. (2021) sought to replicate Bortoloti et al. (2019) but employing emotional pictographs (emojis) in the place of emotional faces. The same general pattern of IRAP effects emerged. The authors concluded that the data provided further support for the DAARRE model prediction that IRAP performances are determined by the relative overlap in Cfunc and Crel properties of the stimuli involved and their coherence with the response options.

Extending the HDML to Other Research Areas

Conceptualizing and interpreting work through the lens of the HDML has also lent itself to other areas within the behavior-analytic literature such as naming, rule-governed behavior and perspective-taking. In the case of naming, recent research has begun to explore variables related to orienting in young children in the context of learning to name stimuli (Sivaraman et al., 2021). Specifically, this study aimed to determine if a child learning to name a novel stimulus involved *deriving* a bidirectional stimulus relation when orienting responses occur simultaneously with the presentation of the spoken name. When the orienting response and the name are presented simultaneously it could be argued that the bidirectional relation between the object and the name is established *directly*, rather than via derivation, because the child hears the name while oriented towards the object, and the child is oriented towards the object while they hear the name. If the orienting response toward the object occurs before the name is presented, and the object is then hidden from view, successful naming by the child would require the *derivation* of a bidirectional relation (object-name leads to a derived name-object relation). Interestingly, the study showed that including a brief delay between the child's orienting response towards a stimulus, and the experimenter naming the stimulus (the stimulus was hidden from view during the delay) undermined the child's ability to point to the stimulus when subsequently asked to do so. Previous research had shown that when young children are oriented towards a stimulus and the name was presented simultaneously, the participants typically identified the stimulus correctly without any additional training. The study thus served to show that the concept of orienting in RFT research could be important in increasing precision in identifying the behavioral processes involved in successful and unsuccessful instances of children learning the names of novel objects.

With respect to research on rule-governed behavior, a relatively new research program has manipulated the levels of derivation, coherence, and flexibility of mutually entailed relations, combinatorially entailed relations, and relational networks, and assessed their impact on persistent derived rule-following (e.g., Bern et al., 2021; Bianchi et al., 2021; Harte & Barnes-Holmes, 2021; Harte et al., 2017, 2018, Harte, Barnes-Holmes, Barnes-Holmes et al., 2020, Harte, Barnes-Holmes, Barnes-Holmes, et al., 2021, Harte, Barnes-Holmes, Moreira et al., 2021), both at the group and single-participant levels. Approaching rule persistence from this perspective has also led to substantive conceptual developments in this area (Harte & Barnes-Holmes, 2021a; Harte, Barnes-Holmes, Barnes-Holmes, & Kissi, 2020). As another example, the interpretation of patterns of IRAP effects in the context of perspective-taking and deictic relational responding has also benefited from these analyses (Kavanagh et al., 2018, 2019) as well as conceptual interpretations of theory of mind tasks (Kavanagh et al., 2020). More recently, conceptual analyses of the symbolic self and altered states of consciousness have also yielded to a HDML analysis (Harte & Barnes-Holmes, 2021c). While exploring this work in any great detail is beyond the scope of the current article, we mention it here to illustrate that the concepts we have presented seem to readily lend themselves to other areas within behavior analysis. Of course, their long-term value as concepts within the literature will depend on whether or not they continue to be of such use.

Concluding Comments on Recent Empirical Developments

We have just provided an overview of recent experimental research that played a key role in driving a number of conceptual developments within the RFT literature. Specifically, creation of the IRAP as a method for capturing "relational framing in flight" led eventually to identifying patterns of behavior on the IRAP that were difficult to explain using existing RFT concepts. In our view, one of the key issues that needed to be addressed was the lack of a formal framework for conceptualizing RFT's units of analysis. The framework that emerged identifies 20 functional-analytic abstractive relational units within the HDML, which may encourage researchers to engage in less frame-centric analyses (i.e., by explicitly identifying five levels of relational development). The HDML also explicitly highlights some of the variables that seem to be involved in the dynamic nature of AARR itself (i.e., coherence, complexity, derivation, and flexibility). In creating the HDML as a framework for conceptualizing AARR, it was deemed important to link this work to a recent increased focus in the RFT literature on the importance of cooperative acts that may be involved in the emergence of AARR itself. In this context, we recognized that the human infant's world presents stimuli that simply orient and evoke responses, under various motivating conditions, while the infant is exposed to the almost continuous flow of vocal stimuli that caregivers provide in their social/cooperative interactions with the infant. Initially, the vocal stimuli will be meaningless to the infant, but gradually they acquire verbal or symbolic meaning (i.e., appropriate behavior-controlling properties), which serves to establish the ROE-M. As the ROE-M emerges out of the weeks, months, and years of behavioral interactions with the physical and social world in which the child resides, the unique nature of human psychological events takes shape. The environment is no longer simply orienting and evoking responses (under motivating conditions) but is synergizing these responses with the 20 units of analysis presented within the HDML. The world is thus gradually transformed for the verbal child, and metaphorically speaking they begin to ROE-M through the world in which relating, orienting, and evoking, under ever-changing motivating conditions, interact in a constant dynamical and nonlinear flow of behavioral events.

Conclusion

The current article has sought to present an overview of recent conceptual and empirical developments within the RFT literature that have emerged largely from the Odysseus research program and its collaborators. In doing so, we did not seek to present a scholarly historical narrative on RFT developments more generally, but rather one perspective on a number of specific recent developments in the theory. We believe these developments in particular are worth presenting because they, in our view, substantively extend the RFT account of language and cognition beyond the 2001 volume. As such, it may be useful for critiques of RFT going forward (be they positive or negative) to address some of the developments we have presented here.

As described above, these developments have emerged largely from work using the IRAP and the need to explain patterns of behavior produced on it. One might argue that it is unwise to use findings from primarily one methodology to build a whole framework (the HDML). And similarly, it may be seen as conceptually impetuous to use the findings of a single study on the impact of pepper sauce on an IRAP performance to introduce motivation as a critical variable in the ROE-M. On balance, this approach in which findings from the basic research laboratory are used to help develop (and constrain) conceptual units of analysis is a hallmark of the inductive, behavior-analytic tradition (see Sidman, 1960). As such, the HDML and ROE-M are not hypothetico-deductive constructs that can be tested in a single study but are suggested units of analysis that will either help to guide and shape future research or not.

We also recognize that many of the concepts we have introduced here may link to, and overlap with, other areas in the experimental psychological literature. For example, relating in the ROE-M could be seen as simply referring to (higher) cognition. 'Orienting' as a concept certainly relates to the vast literature on 'attention', and 'evoking' is clearly relevant to the concept of 'affect'. Furthermore, motivation is directly relevant to concepts that already exist inside behavior analysis (i.e., establishing/motivational operations), and even inside RFT itself (i.e., the concept of a motivative augmental). In addition, specific behavioral effects that we have identified on the IRAP are likely relevant to 'anchoring effects' and/or salience asymmetry effects in binary-based tasks generally (e.g., Bahník et al., 2017; Rothermund & Wentura, 2004). We certainly do not deny all of these and many other potentially important links and connections to the wider psychological literature. Attempting to address, even in some relatively minor or trivial way, all of these connections would simply not be possible in a single article such as the one we have written here. Indeed, a scholarly treatment of how recent developments inside RFT, as articulated herein, sit within the vast literature in behavior analysis, and psychological science generally, would almost certainly require multiple encyclopedic volumes to do the task justice. As noted at the outset, however, our task was necessarily far more humble. We simply aimed to convey, as best we could, a recent 5-year effort to take another small step towards developing a single overarching account of human language and cognition within the behavior-analytic tradition.

References

Bahník, Š., Englich, B., & Strack, F. (2017). Anchoring effect. In R. F. Pohl (Ed.), *Cognitive illusions: Intriguing phenomena in thinking, judgement, and memory* (pp. 223–241). Routledge/Taylor & Francis Group.

Barnes, D., & Holmes, Y. (1991). Radical behaviorism, stimulus equivalence and human cognition. *The Psychological Record*, *41*, 19-30. https://doi.org/10.1007/BF03395091

Barnes-Holmes, D., Barnes-Holmes, Y., Hussey, I., & Luciano, C. (2016). Relational frame theory: Finding its historical and intellectual roots and reflecting upon its future development: Introduction to part II. In R. D. Zettle, S. C. Hayes, D. Barnes-Holmes, & A. Biglan (Eds.), *The Wiley handbook of contextual behavioral science* (pp. 117-128). John Wiley & Sons, Ltd.

Barnes-Holmes, D., Barnes-Holmes, Y., Luciano, C., & McEnteggart, C. (2017). From IRAP and REC model to a multi-dimensional multi-level framework for analysing the dynamics of arbitrarily applicable relational responding. *Journal of Contextual Behavioral Science*, *6*(4), 473-483. 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. *The Psychological Record*, *70*, 605-624. https://doi.org/10.1007/s40732-019-00372-3

Barnes-Holmes, D., Barnes-Holmes, Y., McEnteggart, C., & Harte, C. (2021). Back to the future with an up-dated version of RFT: More field than frame? *Perspectivas em Análise do Comportamento*, *12*(1), 33-51. https://doi.org/10.18761/PAC.2021.v12RFT.03

Barnes-Holmes, D., Barnes-Holmes, Y., Stewart, I., & Boles, S. (2010). A sketch of the implicit relational assessment procedure (IRAP) and the relational elaboration and coherence (REC) model. *The Psychological Record*, *60*, 527-542. https://doi.org/10.1007/BF03395726

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., Hayden, E., Barnes-Holmes, Y., & Stewart, I. (2008). The Implicit Relational Assessment Procedure (IRAP) as a response-time and event-related-potentials methodology for testing natural verbal relations: A preliminary study. *The Psychological Record*, *58*, 497–516. https://doi.org/10.1007/BF03395634

Barnes-Holmes, D., O'Hora, D., Roche, B., Hayes, S. C., Bissett, R. T., & Lyddy, F. (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). Plenum.

Barnes-Holmes, D., & Sivaraman, M. (2020, August 14). Updating RFT: Cooperation came first, the ROE as a unit of analysis, and engineering prosocial behavior. https://science.abainternational.org/up-dating-rft-cooperation-came-first-the-roe-as-a-unit-of-analysis-and-engineering-prosocial-behavior/louise-mchughucd-ie/

Barnes-Holmes, Y. (2001). *Analyzing relational frames: Studying language and cognition in young children* [Unpublished doctoral thesis, National University of Ireland Maynooth].

Baum, W. M. (2018). Three laws of behavior: Allocation, induction, and covariance. *Behavior Analysis: Research and Practice*, *18*, 239-251. https://doi.org/10.1037/bar0000104

Bern, R., Persdotter, T., Harte, C., & Barnes-Holmes, D. (2021). Relational coherence and persistent rule-following: The impact of targeting coherence in a 'non-critical' component of a relational network. *The Psychological Record*, *71*, 279-290. https://doi.org/10.1007/s40732-020-00414-1

Bianchi, P. H, Perez, W. F, Harte, C., & Barnes-Holmes, D. (2021). Effects of coherence and speaker preference on rule-following. *Perspectivas em Análise do Comportamento*, *12*(1), 214-227. https://doi.org/10.18761/PAC.2021.v12.RFT.03

Bortoloti, R., de Almeida, R. V., de Almeida, J. H., & de Rose, J. C. (2019). Emotional faces in symbolic relations: A happiness superiority effect involving the equivalence paradigm. *Frontiers in Psychology*, *10*, 1–12. https://doi.org/10.3389/fpsyg.2019.00954

Bortoloti, R., de Almeida, R.V., de Almeida, J. H., & de Rose, J. C. (2020). A commentary on the dynamics of arbitrarily applicable relational responding involving positively valenced stimuli and its implications for the IRAP research. *The Psychological Record*, *71*, 481-486. https://doi.org/10.1007/s40732-020-00413-2

Bray, E. E., Gnanadesikan, G. E., Horschler, D. J., Levy, K. M., Kennedy, B. S., Famula, T. R., & MacLean, E. L. (2021). Early emerging and highly heritable sensitivity to human communication in dogs. *Current Biology*, *31*, 1-5. https://doi.org/10.1016/j.cub.2021.04.055

Cowie, S. (2018). Behavioral time travel: Control by past, present, and potential events. *Behavior Analysis: Research and Practice*, *18*(2), 174-183. https://doi.org/10.1037/bar0000122

DeCasper, A. J., & Spence, M. J. (1986). Prenatal maternal speech influences newborns' perception of speech sounds. *Infant Behavior*, *9*(2), 133-150. https://doi.org/10.1016/0163-6383(86)90025-1

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. https://doi.org/10.1901/jeab.1986.46-243

Dixon, M. R. (2016). *The PEAK relational training system: Transformation module*. Shawnee Scientific Press.

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. https://doi.org/10.1901/jeab.2007.45-05

Dougher, M. J., Twohig, M. P., & Madden, G. J. (Eds.). (2014). Stimulus-stimulus relations [Special Issue]. *Journal of the Experimental Analysis of Behavior*, *101*(1), 130-170.

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. https://doi.org/10.1901/jeab.2000.73-5

Dymond, S., & Barnes, D. (1995). A transformation of self-discrimination response functions in accordance with the arbitrarily applicable relation of sameness, more-than, and less-than. *Journal of the Experimental Analysis of Behavior, 64*, 163-184. https://doi.org/10.1901/jeab.1995.64-163

Dymond, S., & Roche, B. (2012). *Advances in relational frame theory: Research and application*. New Harbinger.

Finn, M., Barnes-Holmes, D., Hussey, I., & Graddy, J. (2016). Exploring the behavioral dynamics of the implicit relational assessment procedure: The impact of three types of introductory rules. *The Psychological Record, 66*(2), 309–321. https://doi.org/10.1007/s40732-016-0173-4

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

Finn, M., Barnes-Holmes, D., McEnteggart, C., & Kavanagh, D. (2019). Predicting and influencing the single trial-type dominance effect. *The Psychological Record, 69*(3), 425–435. https://doi.org/10.1007/s40732-019-00347-4

Gervain, J. (2015). Plasticity in early language acquisition: The effects of prenatal and early childhood experience. *Current Opinion in Neurobiology, 35*, 13-20. https://doi.org/10.1016/j.conb.2015.05.004

Gomes, C., Perez, W., de Almeida, J., Ribeiro, A., de Rose, J., & Barnes-Holmes, D. (2020). Assessing a derived transformation of functions using the implicit relational assessment procedure under three motivative conditions. *The Psychological Record, 69*, 487–497. https://doi.org/10.1007/s40732-019-00353-6

Greenwald, A. G., McGhee, D. E., & Schwartz, J. L. K. (1998). Measuring individual differences in implicit cognition: The implicit association test. *Journal of Personality and Social Psychology, 74*(6), 1464-1480. https://doi.org/10.1037//0022-3514.74.6.1464

Hare, B., Call, J., & Tomasello, M. (1998). Communication of food location between food and dog (*canis familaris*). *Evolution of Communication, 2*(1), 137–159. https://doi.org/10.1075/eoc.2.1.06har

Harte, C., & Barnes-Holmes, D. (2021a). The status of rule-governed behavior as pliance, tracking and augmenting within relational frame theory: Middle-level rather than technical terms. *The Psychological Record*. Advance Online Publication. https://doi.org/10.1007/s40732-021-00458-x

Harte, C. & Barnes-Holmes, D. (2021b). A primer on relational frame theory (RFT). In M. P. Twohig, M. E. Levin, & J. M. Peterson (Eds.). *The Oxford handbook of acceptance and commitment therapy*. Oxford University Press. https://doi.org/10.1093/oxfordhb/9780197550076.013.4

Harte, C., & Barnes-Holmes, D. (2021c, March 29). Wherever I "ROE-M", there I am: An RFT (technical) analysis of the verbal self and altered states of consciousness. https://science.abainternational.org/wherever-i-roe-m-there-i-am-an-rft-technical-account-of-the-verbal-self-and-altered-states-of-consciousness/louise-mchughucd-ie/

Harte, C., Barnes-Holmes, D., Barnes-Holmes, Y., & Kissi, A. (2020). The study of rule-governed behavior and derived stimulus relations: Bridging the gap. *Perspectives on Behavior Science, 43*, 361-385. https://doi.org/10.1007/s40614-020-00256-w

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. *Behavioral Processes, 157*, 36-46. https://doi.org/10.1016/j.beproc.2018.08.005

Harte, C., Barnes-Holmes, D., Barnes-Holmes, Y., & McEnteggart, C. (2021). Exploring the impact of coherence (through the presence versus absence of feedback) and levels of derivation on persistent rule-following. *Learning and Behavior, 49*, 222-239. https://doi.org/10.3758/s13420-020-00438-1

Harte, C., Barnes-Holmes, D., Barnes-Holmes, Y., McEnteggart, C., Gys, J., & Hassler, C. (2020). Exploring the potential impact of relational coherence on persistent rule-following: The first study. *Learning and Behavior, 48*, 373-391. https://doi.org/10.3758/s13420-019-00399-0

Harte, C., Barnes-Holmes, D., Moreira, M., de Almeida, J. H., Aparecida-Passarelli, D., & de Rose, J. C. (2021). Exploring a Training IRAP as a single participant context for analyzing reversed derived relations and persistent rule-following. *Journal of the Experimental Analysis of Behavior, 115*(2), 460-480. https://doi.org/10.1002/jeab.671

Hayes, S. C. (1986, June). *What is a verbal stimulus?* Invited address presented to the Summer Institute on Verbal Relations, Behaviorists European Summer Academy, Bad Kreuznach, West Germany.

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

Hayes, S. C., & Brownstein, A. J. (1985, May). *Verbal behavior, equivalence classes, and rules: New definitions, data, and directions*. Invited address presented at the meeting of the Association for Behavior Analysis, Columbus, OH.

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 language and cognition. *Journal of the Experimental Analysis of Behavior, 101*(1), 112–129. https://doi.org/10.1002/jeab.64

Horne, P. J., & Lowe, C. F. (1996). On the origins of naming and other symbolic behavior. *Journal of the*

Experimental Analysis of Behavior, *65*, 185-241. https://doi.org/10.1901/jeab.1996.65-185

Hughes, S., & Barnes-Holmes, D. (2016a). Relational frame theory: The basic account. In R. D. Zettle, S. C. Hayes, D. Barnes-Holmes, & A. Biglan (Eds.), *The Wiley handbook of contextual behavioral science* (pp. 129-178). John Wiley & Sons, Ltd.

Hughes, S., & Barnes-Holmes, D. (2016b). Relational frame theory: Implications for the study of human language and cognition. In R. D. Zettle, S. C. Hayes, D. Barnes-Holmes, & A. Biglan (Eds.), *The Wiley handbook of contextual behavioral science* (pp. 129-178). John Wiley & Sons, Ltd.

Kavanagh, D., Barnes-Holmes, Y., & Barnes-Holmes, D. (2020). The study of perspective taking: Contributions from mainstream psychology and behavior analysis. *The Psychological Record*, *70*, 581-604. https://doi.org/10.1007/s40732-019-00356-3

Kavanagh, D., Barnes-Holmes, Y., Barnes-Holmes, D., McEnteggart, C., & Finn, M. (2018). Exploring differential trial-type effects and the impact of a read-aloud procedure on deictic relational responding on the IRAP. *The Psychological Record*, *68*, 163–176. https://doi.org/10.1007/S40732-018-0276-1.

Kavanagh, D., Roelandt, A., Van Raemdonck, L., Barnes-Holmes, Y., Barnes-Holmes, D., & McEnteggart, C. (2019). The on-going search for perspective-taking IRAPs: Exploring the potential of the natural language IRAP. *The Psychological Record*, *69*(2), 291–314. https://doi.org/10.1007/s40732-019-00333-w

Keuleers, E., Diependaele, K., & Brysbaert, M. (2010). Practice effects in large-scale visual word recognition studies: A lexical decision study on 14,000 Dutch mono- and di-syllabic words and nonwords. *Frontiers in Psychology*, *1*, 174. https://doi.org/10.3389/fpsyg.2010.00174

Leech, A., Barnes-Holmes, D., & Madden, L. (2016). The implicit relational assessment procedure (IRAP) as a measure of spider fear, avoidance, and approach. *The Psychological Record*, *66*, 337–349. https://doi.org/10.1007/s40732-016-0176-1

Leech, A., Barnes-Holmes, D., & McEnteggart, C. (2017). Spider fear and avoidance: A preliminary study of the impact of two verbal rehearsal tasks on a behavior–behavior relation and its implications for an experimental analysis of defusion. *The Psychological Record*, *67*, 387–398. https://doi.org/10.1007/s40732-017-0230-7

Lipkens, R., Hayes, S. C., & Hayes, L. J. (1993). Longitudinal study of derived stimulus relations in an infant. *Journal of Experimental Child Psychology*, *56*, 201–239. https://doi.org/10.1006/jecp.1993.1032

Lowenkron, B. (1991). Joint control and the generalization of selection-based verbal behavior. *The Analysis of Verbal Behavior*, *9*, 121–126. https://doi.org/10.1007/BF03392866

Maloney, E., & Barnes-Holmes, D. (2016). Exploring the behavioral dynamics of the implicit relational assessment procedure: The role of relational contextual cues versus relational coherence indicators as response options. *The Psychological Record*, *66*, 395–403. https://doi.org/10.1007/s40732-016-0180-5

McHugh, L., Barnes-Holmes, Y., & Barnes-Holmes, D. (2004). Perspective-taking as relational responding: A developmental profile. *The Psychological Record*, *54*, 115- 144. https://doi.org/10.1007/BF03395465

Michael, J. (1993). Establishing operations. *The Behavior Analyst*, *16*(2), 191–206. https://doi.org/10.1007/BF03392623

Michael, J. (2007). Motivating operations. In O. J. Cooper, T. E. Heron, & W. L. Heward (Eds.), *Applied behavior analysis* (2nd ed., pp. 374–391). Merrill Prentice Hall.

Moon, C., Cooper, R. P., & Fifer, W. P. (1993). Two-day-olds prefer their native language. *Infant Behavior & Development*, *16*(4), 495–500. https://doi.org/10.1016/0163-6383(93)80007-U

Moon, C., Lagercrantz, H., & Kuhl, P. K. (2013). Language experienced in utero affects vowel perception after birth: A two-country study. *Acta Paediatrica*, *102*(2), 156-160. https://doi.org/10.1111/j.1467-7687.2010.00978.x

Nickerson, R. S. (1998). Confirmation bias: A ubiquitous phenomenon in many guises. *Review of General Psychology*, *2*(2), 175–220. https://doi.org/10.1037/1089-2680.2.2.175

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. 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

O'Shea, B. A., Watson, D. G., & Brown, G. (2016). Measuring implicit attitudes: A positive framing bias flaw in the implicit relational assessment procedure (IRAP). *Psychological Assessment*, *28*(2), 159–170. https://doi.org/10.1037/pas0000172

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. https://doi.org/10.1007/BF03395652

Pinto, J. A. R., de Almeida, R. V., & Bortoloti, R. (2020). The stimulus' orienting function may play an important role in IRAP performance: Supportive evidence from an eye-tracking study of brands. *The Psychological Record*, *70*, 257–266. https://doi.org/10.1007/s40732-020-00378-2

Rothermund, K., & Wentura, D. (2004). Underlying processes in the Implicit Association Test: Dissociating salience from associations. *Journal of Experimental Psychology: General*, *113*, 139-165. https://doi.org/10.1037/0096-3445.133.2.139

Ruiz, F. J., & Luciano, C. (2011). Cross-domain analogies as relating derived relations among two separate relational networks. *Journal of the Experimental Analysis of Behavior*, *95*, 369–385. https://doi.org/10.1901/jeab.2011.95-369

Schmidt, M., de Rose, J. C, & Bortoloti, R. (2021). Relating, orienting and evoking functions in an IRAP study involving emotional pictographs (emojis) used in electronic messages. *Journal of Contextual Behavioral Science*, *21*, 80-87. https://doi.org/10.1016/j.jcbs.2021.06.005

Sidman, M. (1960). *Tactics of scientific research: Evaluating experimental data in psychology*. Basic Books.

Sidman, M. (1971). Reading and auditory-visual equivalences. *Journal of Speech, Language, and Hearing Research*, *14*, 5-13. https://doi.org/10.1044/jshr.1401.05

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

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*, 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), 519–555. https://doi.org/10.1002/jeab.715

Skinner, B. F. (1953). *Science and human behavior*. Macmillan.

Skinner, B. F. (1957) *Verbal behavior*. Appelton-Century-Crofts.

Steele, D. L., & Hayes, S. C. (1991). Stimulus equivalence and arbitrarily applicable relational responding. *Journal of the Experimental Analysis of Behavior*, *56*, 519–555. https://doi.org/10.1901/jeab.1991.56-519

Stewart, I., & Barnes-Holmes, D. (2004). Relational frame theory and analogical reasoning: Empirical investigations. *International Journal of Psychology and Psychological Therapy*, *4*, 241–262.

Vahey, N., Nicholson, E., & Barnes-Holmes, D. (2015). A meta-analysis of criterion effects for the implicit relational assessment procedure (IRAP) in the clinical domain. *Journal of Behavior Therapy and Experimental Psychiatry*, *48*, 59–65. https://doi.org/10.1016/j.jbtep.2015.01.004.

Villatte, M., Villatte, J., & Hayes, S. C. (2016). *Mastering the clinical conversation: Language as intervention*. Guilford.

Wilson, D. S. (2007). *Evolution for everyone: How Darwin's theory can change the way we think about our lives*. Delacorte Press.

Wilson, D. S., Hayes, S. C., Biglan, A., & Embry, D. D. (2014). Evolving the future: Toward a science of intentional change. *Behavioral and Brain Sciences*, *37*(4), 395-416. https://doi.org/10.1017/S0140525X13001593

Zapparoli, H. B., Marin, R., & Harte, C. (2021). Rule-governed behavior: An ongoing RFT-based operant analysis. *Perspectivas em Análise do Comportamento*, *12*(1), 197-213. https://doi.org/10.18761/PAC.2021.v12.RFT.09

Zettle, R. D., Hayes, S. C., Barnes-Holmes, D., & Biglan, A. (2016). *The Wiley handbook of contextual behavioral science*. John Wiley & Sons, Ltd.

Received: July 21, 2021
Final Acceptance: December 11, 2021
Editor-in-Chief: Mark Galizio
Associate Editor: Mark Galizio