perspectivas
em análise do comportamento
ISSN 2177-3548
centro paradigma ciências do comportamento
www.revistaperspectivas.org

# Rule-governed behavior: An ongoing RFT-based operant analysis

## Comportamento governado por regras: Uma análise operante baseada na RFT

## Conducta gobernada por reglas: un análisis operante basado en la RFT

**Heloisa Ribeiro Zapparoli**[1], **Ramon Marin**[1], **Colin Harte**[2]

[1] Federal University of São Carlos, São Carlos, São Paulo, Brazil (Universidade Federal de São Carlos) [2] Department of Psychology, National College of Ireland, Dublin, Ireland | **Título abreviado**: RFT-based operant analysis of rules | **Endereço para correspondência:** Portuguese version: Universidade Federal de São Carlos - Rodovia Washington Luiz, km 235, CEP: 13565-905 - São Carlos, SP, Brasill **Email**: Portuguese version: helozapparoli@gmail.com or rm.ramonmarin@gmail.com; English version: colin.n.harte@gmail.com | **doi**: 10.18761/PAC.2021.v12.RFT.09

**Abstract:** Rule-governed behavior is broadly defined as verbal antecedent stimuli that specify dependence relations between stimuli and events. Since its conception, this definition has supported a relatively rich program of research within the experimental analysis of behavior. Specifically, researchers have sought to explore the extent to which verbal rules are involved in operant behavior, both in the basic and applied domains. However, some have highlighted the need for a more complete understanding of what "*specification*" means in the context of rule-following and behavior analysis. The current article aims to present an operant account of what it means to understand and follow verbal rules, drawing largely on stimulus equivalence, and focusing in particular on a relational frame theory (RFT) perspective. To this end, we provide an overview of an RFT-based operant account of rule-following as it currently stands, and outline a recent program of experimental research that has utilized this approach to explore the complexities involved in rule-following in the face of competing reinforcement contingencies, a phenomenon typically linked to human psychological suffering. Implications for going forward in developing a more complete operant account of rule-governed behavior in both the basic and applied domains are considered.

**Keywords:** rules; rule-following; operant behavior; relational frame theory; derived relations.

The authors would like to thank Dermot Barnes-Holmes for his invaluable comments and feedback on earlier versions of this manuscript. The first and second authors were supported by the Brazilian Coordination for the Improvement of Higher Education Personnel (CAPES; Finance Code 001).

**Resumo:** O comportamento governado por regras é amplamente definido como estímulos antecedentes verbais que especificam relações de dependência entre estímulos e eventos. Desde sua formulação, essa definição tem dado suporte a um programa relativamente rico de pesquisas experimentais dentro da análise experimental do comportamento. Especificamente, pesquisadores têm buscado explorar até que ponto as regras verbais estão envolvidas no comportamento operante, nos domínios da pesquisa básica e aplicada. No entanto, alguns deles destacaram a necessidade de uma compreensão mais ampla do que "especificação" significa no contexto do seguimento de regras e da análise do comportamento. O presente artigo tem como objetivo apresentar uma explicação operante do que significa compreender e seguir regras verbais, baseando-se amplamente na equivalência de estímulos e focando, em particular, na perspectiva da Teoria das Molduras Relacionais (do inglês, *Relational Frame Theory* - RFT). Para isso, fornecemos uma visão atual e geral de uma explicação operante baseada na RFT sobre o seguimento de regras, e apresentamos um programa recente de pesquisa experimental que utilizou esta abordagem para explorar as complexidades envolvidas no seguimento de regras em face de contingências de reforçamento concorrentes, um fenômeno tipicamente relacionado ao sofrimento humano. São consideradas implicações para o avanço do desenvolvimento de uma explicação operante mais completa do comportamento governado por regras, nos domínios da pesquisa básica e aplicada.

**Palavras-chave:** regras; seguimento de regras; comportamento operante; Teoria das Molduras Relacionais; relações derivadas

**Resumen:** Conducta gobernada por reglas es ampliamente definida como estímulos antecedentes verbales que especifican relaciones de dependencia entre estímulos y eventos. Desde su formulación, esta definición ha apoyado un programa de investigación relativamente rico dentro del análisis experimental de la conducta. Específicamente, los investigadores han buscado explorar hasta qué punto las reglas verbales están involucradas en la conducta operante, tanto en los dominios básicos como aplicados. Sin embargo, algunos han destacado la necesidad de una comprensión más completa de lo que significa "especificación" en el contexto del seguimiento de reglas y el análisis de la conducta. Este artículo tiene como objetivo presentar una explicación operante de lo que significa comprender y seguir reglas verbales, basándose en gran medida en la equivalencia de estímulos y centrándose en particular en una perspectiva de la Teoría de los Marcos Relacionales (del inglés, Relational Frame Theory - RFT). Para eso, proporcionamos una descripción general de una explicación operante basada en RFT del seguimiento de reglas tal como está actualmente, y presentamos un programa reciente de investigación experimental que ha utilizado este enfoque para explorar las complejidades involucradas en el seguimiento de reglas frente a cadenas concurrentes de reforzamiento, un fenómeno típicamente relacionado al sufrimiento psicológico humano. Se consideran las implicaciones para avanzar en el desarrollo de una explicación operante más completa de la conducta gobernada por reglas en los dominios de la investigación básica y aplicada.

**Palabras clave:** reglas; seguimiento de reglas; conducta operante; Teoría de los Marcos Relacionales; relaciones derivadas

Rule-governed behavior (or instructional control) has long been recognized as an important concept within the behavior-analytic literature. Skinner (1966) defined rules — or instructions — as antecedent verbal stimuli that specify dependence relations between stimuli and events. Some authors have argued that rule-governed behavior constitutes a key behavioral difference between humans and non-humans. Specifically, it has been argued that the behavior of non-human animals is governed solely by contact with direct contingencies in the environment, while the behavior of humans may also be under the control of verbal stimuli, in particular rules and instructions (Baron et al., 1985; Hayes et al., 1989; Skinner, 1957, Törneke et al., 2008). For example, it would be fruitless to instruct a dog saying "If you bite the guest, you will be punished!" unless specific "guest-friendly" behavioral patterns had been previously established for these terms through direct interventions (e.g., shaping). In contrast, one would expect that a typically developing and verbally-able child would be able to understand and successfully follow a similar rule without such direct intervention (see Paracampo & Albuquerque, 2005 for an introduction).

Rule-following behavior seemingly constitutes a highly adaptive human ability because it allows the individual to learn without having to directly contact environmental contingencies. For instance, for a verbally-able individual, the simple rule "Do not pet the dog, he is aggressive" allows the individual to avoid interacting with the dog without directly contacting negative contingencies (e.g., receiving a nasty bite). Indeed, in some instances, following a simple rule may protect us from life-threatening events. However, despite its benefits, under certain conditions, rule-governed behavior may undermine potentially positive contact with environmental contingencies (Zettle & Hayes, 1982). For example, in the case of the rule about the aggressive dog above, rigid and overgeneralized rule-following may undermine behavioral variability in the presence of other dogs (e.g., the person may come to always avoid all dogs, which could be seen as removing a potential source of reinforcement). Indeed, behavioral variability in and of itself has been suggested as an important tool for adapting to constantly changing environments (see Hayes et al., 1989, for the first book length treatment of the subject). Thus, a decrease in variability as a result of excessive and overgeneralized rule-following, by definition, may undermine sensitivity to new direct contingencies of reinforcement and has been argued to form the basis for some forms of psychological distress (e.g., Generalized Anxiety — Friman et al. 1998; Luciano et al., 2004; Törneke et al., 2008). Indeed, this general idea has been one of the core tenets of a prominent third-wave behavior therapy[1], acceptance and commitment therapy (ACT; Hayes et al., 1999).

Some researchers have thus highlighted the need for a more complete technical account of the basic behavioral processes involved in rule-following behavior (Harte, Barnes-Holmes, Barnes-Holmes, & Kissi, 2020; Barnes-Holmes et al., 2001; Törneke et al., 2008). Specifically, these researchers have argued that the definition of rules originally provided by Skinner (1966) did not sufficiently explain how rules actually *specify* contingencies. If 'specify' is simply seen as referring to discriminative stimuli, it remains unclear how a rule may control behavior in the absence of direct contact with contingencies; discriminative stimuli, by definition, acquire their controlling properties through direct contact with differential reinforcement contingencies (Catania, 1979, 1999).

Progress on providing a functional-analytic definition of specification was made by Murray Sidman and his colleagues with their work on equivalence relations, which first emerged only five years after Skinner's seminal work on instructional control (see Sidman, 1994, or de Rose et al., 2014, for book-

---

[1] The term 'third-wave' has typically come to be employed in distinguishing different generations of behavior therapies in terms of their changing therapeutic focus. Specifically, relative to their second-wave counterparts, third-wave behavior therapies generally tend to deemphasise attempts to alter the cognitive content of the client and instead focus on changing the client's relationship with this cognitive content (see Hayes, 2004, for a detailed description). Apart from ACT, other prominent third-wave behavior therapies include dialectical behavior therapy (DBT; Linehan, 1993), mindfulness-based cognitive therapy (MBCT; Segal et al., 2002), and functional analytic psychotherapy (FAP; Kohlenberg & Tsai, 1991). See Lucena-Santos et al., 2015, and Leonardi, 2015, for further discussion of the history of different generations of behavior therapies.

length reviews in English and Portuguese respectively). The basic stimulus equivalence effect was first identified in the context of teaching basic reading skills to individuals with learning disabilities. A critical and unexpected finding that emerged from this work was that after training a small number of conditional relations between stimuli (i.e., between spoken words and pictures, and between spoken words and written words), a number of untrained relations emerged (i.e., between pictures and written words and between written words and pictures; Sidman, 1971). That is, humans were capable of deriving relations between stimuli, in the absence of direct reinforcement (see de Rose, 2012; de Rose et al., 2012, for related research of this nature and de Rose & Almeida, 2019, for an introduction).

For illustrative purposes, imagine the simple instruction given to a child "That red berry will make you very sick". If particular words in this rule participate in equivalence relations with the actual stimuli (e.g., the words "red" and "berry" with the actual colour red and the object berry), then the rule may be followed, in part, because of these derived equivalence relations. Of course, rules involve more than just equivalence relations. Typically, a rule is seen as specifying a contingency, which involves establishing some sort of conditional relation between or among events — in this case between eating the red berry and being sick. The conditionality seemingly requires more than basic equivalence relations between words and objects or events. Additional types or classes of relational responses are required. The emergence of what came to be known as relational frame theory (RFT; see Hayes et al., 2001, for the seminal full book-length treatment) appeared to provide this additional step. Indeed, the very origins of RFT itself commenced with an effort to explain rule-governed behavior within a behavioral and functional-analytic framework (Hayes & Hayes, 1989).

In arguing that equivalence relations, combined with other types of relations such as conditionality, provide a behavior-analytic account of rule-governed behavior it is necessary that the account explains derived relational responding itself. Simply pointing to the ability to form equivalence relations, for instance, does not constitute a satisfactory explanation for symbolic relations because equivalence responding itself remains unexplained, at least in behavioral terms (Hayes, 1991; Steele & Hayes, 1991). For this reason, RFT not only argued that there are multiple types of derived relations, but these classes were essentially established as generalized operants, through interactions with the wider verbal community. In explaining how young children learn to form equivalence relations, RFT argued that the early verbal interactions involved in naming were likely very important. Imagine, for example, a child playing with a teddy bear in the presence of a parent saying "Look at the teddy bear", and directly reinforcing the child for pointing or looking at the bear. Across multiple occurrences, a relationship between the toy and the spoken word /*teddy bear*/ may be established. In addition, the child is also directly reinforced for producing other appropriate responses such as pointing to the bear upon the parent asking "Where is the teddy bear?", and for saying "Teddy" when the bear is observed or in response to other appropriate contextual cues such as "What is this?". Across multiple exemplars of coordinating many different objects, situations, locations, etc. with their names in multiple other contexts, the operant class of equivalence (or coordination in RFT terms) comes to be established for novel exemplars presented in similar contexts without further need for direct training or reinforcement (once that particular generalized operant is established sufficiently in the individual's behavioral history). For instance, if someone subsequently introduces themself to the child by saying "Hello, I'm Steven!", when asked "Where is Steven?", the child may respond by pointing to that person in the absence of a history of direct reinforcement for doing so.

As mentioned above, RFT argues that other relations in addition to coordination or equivalence are likely required when it comes to rule following. Basic temporal relations, such as *before-after* and slightly more advanced types, such as *if-then conditionality,* would appear to be essential for even the most basic rule-governed behavior. According to RFT these other types of relations are established in a broadly similar manner to equivalence relations. Specifically, interactions with the wider verbal community serve to establish specific contextual cues that control appropriate relational responses in accordance with these relations. For example, operant contingencies establish responses to new cues such as "if", "then", "before" and "after". Again, across numerous multiple exemplars, the generalized operant class of conditionality (i.e., "if X then Y" derives "Y depends on X") comes to be established without further need for direct training or reinforcement. Thus, when a child hears "if you do x, y will happen", or in more concrete terms "if you eat this berry you will be sick" the child "understands" the rule and may thus behave accordingly (i.e., does not eat the berry). According to RFT, therefore, multiple-exemplars provide the overarching or generalized operant contingencies that serve to establish what are called relational frames and increasingly complex networks that in essence provide a behavior-analytic account of instructional control or rule-governed behavior.

## RFT: Some more details

We have thus far provided a general overview of how RFT conceptualizes rules as combinations of relations into networks of relational frames. Before continuing, it seems wise to provide some additional details on the key terms and concepts of RFT itself, so that these can be referenced later in the current article.

One critical feature of RFT that clearly extends beyond the language of equivalence relations is that more generic terms are required for the key properties of a relational frame (as opposed to an equivalence relation). Unlike the equivalence class or relation, which is defined as symmetry (if a=b then b=a) and transitivity (if a=b and a=c, then b=c and c=b), RFT invokes the concepts of mutual and combinatorial entailment. Similar to symmetry, mutual entailment refers to the specific bidirectional derived relation between two stimuli, and like transitivity, combinatorial entailment refers to relations that can be derived between and among stimuli when three or more elements are related. The key difference within RFT, however, is that the relating behaviour observed in emergent equivalence responding is considered to be one class of generalized operant behaviour, but many other classes are possible (i.e., relational frames; Steele & Hayes, 1991; see Perez et al., 2013, for an introduction in Portuguese). As mentioned above, during the course of early language learning human children are taught to respond in accordance with relational frames, such as temporality, conditionality, opposition, difference, comparison (e.g., bigger versus smaller than), and so on. Thus, within RFT, the more generic terms of mutual and combinatorial entailment are employed to account for the fact that derived relations other than equivalence are possible and are likely involved in increasingly complex forms of relating, such as rule-following. For example, if A is greater than B, this entails that B is less than A, not B is greater than A (i.e., the relation is not symmetrical). Or in more concrete terms, consider the simple rule mentioned above, which stated "if you eat that berry you will get sick". This relation does not necessarily entail "if you get sick then you ate the berry" (again, the relation is not symmetrical -- you may get sick for some other reason). Thus, more generic terms than symmetry (i.e., mutual entailment) and transitivity (i.e., combinatorial entailment) are invoked.

In addition, RFT also invokes the concept of the transformation of stimulus functions which describes the way in which functions of stimuli are changed in accordance with the relation operating between them, in the absence of direct reinforcement (Dymond & Rehfeldt, 2000). For RFT, this property is key to accounting for the ways in which stimuli or events come to acquire, change, or lose psychological properties. The term 'transformation' is employed within RFT to once again account for the fact that the functions of stimuli participating in relations other than equivalence transform in accordance with the specific relation involved. That is, the same function does not necessarily emerge among all participating stimuli - the nature of the transformation of stimulus functions depends on the specific relations involved (e.g., Dymond & Barnes, 1995). For example, imagine a situation in which a child has been bitten by a relatively small dog and later learns that a neighbor owns a very large dog. Based on the transformation of fear functions, in accordance with the comparative relation between the two dogs (in this case, smaller/larger), it is possible that the neighbor's larger dog will evoke even greater fear and avoidance than the smaller dog that actually bit the child in the first place (see Dougher et al., 2007 for relevant experimental evidence).

Finally, in making a distinction between entailment and transformations of functions, RFT specifies that these properties operate under separate classes of contextual control. Specifically, the type of relation between stimuli is determined by *Crel* contextual cues (e.g., if/then, before/after, etc.), while the functions produced during the act of relating are determined by *Cfunc* contextual cues (e.g., reinforcing, discriminative, eliciting, etc.). Thus, for example, in the sentence "John is the opposite of Mark in sports", the term "is the opposite of" operates as a Crel, because it establishes the type of relationship between John and Mark, and "in sports" operates as a Cfunc, because it establishes a particular type of function in this relation (Torneke, 2010). According to RFT, relational frames may combine into increasingly complex relational networks, which helps to explain scaling up to more complex levels of human language and cognition, one of which is rule-following to which we now turn.

## Rules as derived relational networks within RFT

As explained previously, a rule or instruction can be thought of as a network of relational frames, typically involving coordination and other types of relations under the control of appropriate Crels. In addition, rules sometimes contain explicit Cfuncs that help to transform the behavioral control functions for stimuli specified within the network (Barnes-Holmes et al., 2001). For example, consider a person who starts to approach a friend's dogs, and receives the instruction "My dogs are very different in temperament, so feel free to pet one but not the other." As explained previously, the rule involves frames of coordination between the words in the sentence and the actual stimuli and events specified in the rule. However, the critical contextual cues in controlling behavior with respect to the dogs in this case is the Crel "different" and the Cfunc "temperament". That is, the rule informs the listener not to treat the dogs equally in terms of relative friendliness or safety. As an aside, in many cases, rules may not contain a word that can be identified as an isolated Cfunc because both the Crel and Cfunc properties are inherent in the same words. For example, in the instruction, "If you try to pet my dog he will bite you" the word "bite" is coordinate with actual biting (a Crel property) and the psychological impact of the word establishes an avoidance function for the dog (a Cfunc property).

At this point, it should be possible to appreciate how the RFT approach to rule-governed behavior attempts to maintain and develop a generic operant approach to this type of behavior. In effect, operant contingencies in the wider verbal community establish the basic units of reference or specification, and then these units combine into networks thus scaling up into increasingly complex networks that may then function as rules that are spoken and understood within that verbal community. As a result, speakers and listeners interact less and less with direct contingencies and more and more with the rules that specify such contingencies. And thus, as Skinner argued, humans may solve problems by following rules, rather than simply relying upon nothing but trial-and-error type learning.[2]

**Experimental demonstration research**. The foregoing material provides an RFT conceptual analysis of rule-governed behavior in terms of relational networks, but of course the account requires, at minimum, empirical evidence to support the interpretation. At the time of writing two studies had been published which specifically set out to provide this evidence. The first of these was reported by O'Hora et al. (2004). In this study, participants first learned to respond to abstract stimuli as func-

---

[2] The reader should note that we are not here arguing that rules never interact with direct contingencies. Indeed, much other research in the literature to date has highlighted numerous variables related to direct contingencies that seem to modulate rule-following (e.g., Albuquerque et al., 2003; Albuquerque et al., 2006; Cortez & Reis, 2008; de Almeida et al., 2020; Paracampo et al., 2007; Paracampo et al., 2013; Reis et al., 2010; Santos et al. 2004). As we will later argue, it is this very interaction that may be of particular importance in coming to understand the complexities involved in persistent rule-following in the face of competing reinforcement contingencies (or contingency insensitivity) and its implications for psychological distress.

tionally equivalent to the words: “before”, “after”, “same”, and “different”. In effect, participants were first exposed to exemplar training designed to establish four stimuli as contextual cues (in this case Crels) that could then be used to “create” a relational network that would function as a rule for emitting a wide range of sequence responses (see Figure 1 for a summary overview of the procedure employed in the study).

Establishing cues for “before” and “after” involved presenting two shapes on the screen in a particular order (e.g., circle before square) with two sets of stimuli on the bottom left and right corners of the screen which included these stimuli and relational cues (e.g., circle before square or square after circle; see left-hand panel of Figure 1). Reading from bottom to top, the sets of stimuli consisted of two arbitrary shapes (e.g., a circle and a square) with an arbitrary contextual cue for “before” or “after” (e.g., ‘circle ()() square’ and ‘square ()() circle’). Participants were reinforced for choosing one of the two stimulus sets contingent upon the order in which the two sample shapes were presented in the center of the screen (e.g., a circle followed by a square). For example, if a circle was presented in the center of the screen, followed by a square, then choosing ‘circle ()() square’ was reinforced. Thus, the arbitrary contextual cue ‘()()’ acquired the function of “before”. Similarly, if a circle was presented in the center of the screen, followed by a square, then choosing ‘square :::: circle’ was reinforced, thus establishing the function of “after” for a second arbitrary contextual cue (i.e., ‘::::’).

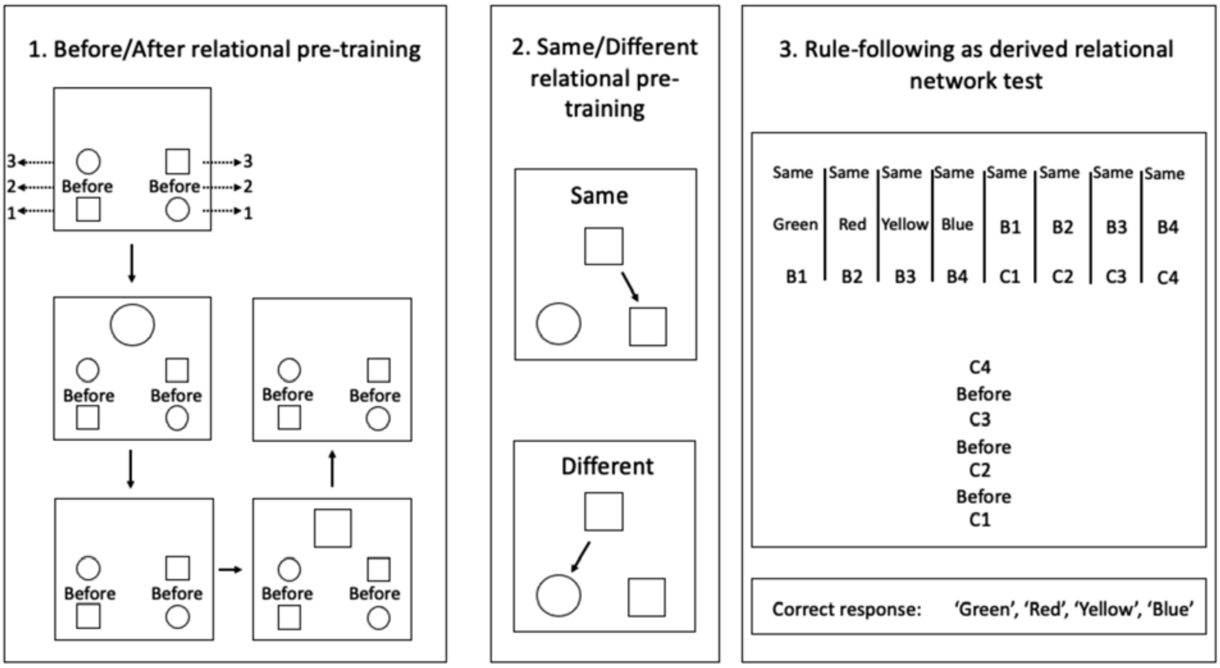

**Figure 1**. An illustration of the experimental training and testing procedures from O’Hora et al. (2004)

**Note**. In the left-hand panel, the numbers 1, 2, and 3, illustrate the order in which these stimuli appeared to participants on the screen. The reader should also note that at no point during pre-training or testing did the words ‘same’, ‘different’, and ‘before’ appear on the screen, but rather the arbitrary stimuli ‘%%%’, ‘!!!’, and ‘()()’, respectively (‘::::’ was presented in the case of ‘after’ pretraining). Relatedly, during the test phase (right-hand panel), the written color words at the top of the screen were presented as squares of the actual colors they describe, while the same words at the bottom of this panel denote the color and sequence of keys participants had to press to emit a correct response in this instance. Finally, B1-B4 and C1-C4 denote novel nonsense syllables (i.e., LIB, DAX, MIM, VEK, CUG, GAN, JOM, MUB) that had not been employed in previous pre-training phases.

Establishing cues for "same" and "different" involved first presenting one of two arbitrary stimuli (%%% or !!!) at the top of a computer screen with a sample shape in the centre of the screen (e.g., a square or a circle), and two comparison shapes at the bottom of the screen (e.g., a square and a circle; see center panel of Figure 1). In the presence of the stimulus that was trained to be functionally equivalent to "same" participants were reinforced for choosing the comparison shape that matched the sample shape in the center of the screen (e.g., square in the presence of a square). In contrast, in the presence of the stimulus that was trained to be functionally equivalent to "different" participants were reinforced for choosing the comparison shape that was different to the sample shape in the center of the screen (e.g., square in the presence of a circle). Upon achieving a particular mastery criterion, participants were exposed to a testing procedure within which no reinforcement was provided, and the samples and comparisons were novel stimuli. Participants were again required to reach a particular mastery criterion to demonstrate that the "same" and "different" contextual cues had been established.

The novel contextual cues for "same", "different", "before" and "after" were then presented with sets of novel nonsense syllables and colored squares (the latter corresponded with colored response keys on the computer keyboard; see right-hand panel of Figure 1). These stimuli were presented in such a manner that it allowed participants to derive if relations between each of the colored squares and the nonsense syllables were "same" or "different". As an illustrative example, the reader is directed to the right-hand panel of Figure 1. The sequence presented along the top of this panel involves establishing 'green' as the same as novel stimulus B1, 'red' as the same as novel stimulus B2, 'yellow' as the same as novel stimulus B3, and 'blue' as the same as novel stimulus B4. Moving over to the right of this sequence, each of those novel B stimuli is then established as the same as four novel C stimuli. That is, B1 as the same as C1, B2 as the same as C2, B3 as the same as C3, and B4 as the same as C4. Establishing that these novel C stimuli meant the same as the novel B stimuli meant that participants should derive (through combinatorial entailment) that the C stimuli meant the same as the colors (e.g., green is the same as B1 which is the same as C1, therefore C1 is the same as green, etc.). The screen also presented an "instruction" using the "before" and "after" cues that specified the sequence in which the colored keys on the keyboard should be pressed. To continue with the illustrative example presented in the right-hand panel of Figure 1, the reader is directed to stimuli presented vertically on the bottom half of this panel. Presented to participants to be read from bottom to top, the sequence C1 before C2 before C3 before C4 is presented. Having correctly derived the meaning of these C stimuli above, correct responding in accordance with this instruction would involve pressing the buttons in the sequence 'green' before 'red' before 'yellow' before 'blue'. In more precise RFT terms, the researchers established relational networks that were shown to control specific response sequences.

In a follow-up study, O'Hora et al. (2014) replicated and extended the basic effect by showing that the derived sequence responding reported in the earlier study could itself be brought under contextual control. Specifically, in the first of two experiments, in one block of trials participants received punishing (i.e., "Wrong") feedback for correctly responding in accordance with the same derived instruction sequences as before, and reinforcing feedback (i.e., "Correct") for responding in any other way. In a subsequent block, participants now received reinforcing feedback for correctly responding in accordance with the derived instruction sequence, and punishing feedback for responding in any other way. In between each of these feedback blocks, participants received a block of trials within which no feedback was provided on responding. Results demonstrated that responding was indeed sensitive to differential consequences. In the second experiment, reinforcing and punishing consequences were varied systematically in the presence of two novel antecedent stimuli, and antecedent control was observed for all participants. Overall, the research provided basic experimental support for a technical RFT analysis of rules as derived relational networks, while also serving to highlight the highly complex relational phenomena that appears to be involved in rule-governed behavior.

**Experimental analyses of rule-governed behavior as relational networking**. The basic experimental demonstration research outlined above certainly provided evidence to support the RFT conceptual analyses of rule-governed behavior. However, much more seems to be required if RFT is to contribute towards a better understanding of the extent to which rules (or more precisely relational networks) actually control behavior in any given instance (see Harte & Barnes-Holmes, 2021, for an extended discussion). More informally, the ability to "understand" a rule is of course critical in generating a behavior-analytic account of such behavior. However, a more complete experimental analysis also requires that we conduct empirical research to identify the critical variables involved in determining if and when rules are actually followed when they have been understood. Indeed, such research would be essential in terms of connecting basic to applied research in this area. For example, an individual suffering from chronic pain may well understand the rule "Regular exercise, even when you are experiencing pain, may reduce or eradicate your pain." However, even if this rule is understood it may fail to produce the desired behavior (i.e., a client may still avoid taking regular exercise). Basic lab-based experimental research should therefore begin to examine the variables that increase or decrease rule-following (as well as rule-understanding). Recently, a program of research has emerged that attempts to address this very issue -- a brief overview of this work is presented subsequently.

The basic preparation involved in the research focused in particular on what is described as persistent rule-following. Specifically, this recent work has attempted to explore the properties or variables involved in derived rule-following in the face of competing reinforcement contingencies. In an initial study, Harte et al. (2017) sought to explore the impact of providing participants with a "direct" rule versus a rule that involved a novel, within-experiment derivation, on persistent rule-following. Participants were either given a rule that specified in natural language how to respond on a subsequent MTS task (i.e., choose the comparison image at the bottom of the screen that is *least like* the sample image at the top of the screen; the direct rule), or a rule that involved a novel within-experiment derivation (the "derived" rule). For the latter group, participants were first required to derive a novel relation between the key phrase *least like* and an unknown word. Specifically, participants were first told that 'least like' is the same as an Irish word 'eagsula', 'eagsula' is the opposite of the welsh word 'un', and that 'un' is the opposite of the Sudanese word 'beda' (see top panel of Figure 2 in which this sequence is presented as A=B-C-D, D=A). They were then asked what 'beda' meant. Upon successfully deriving that 'beda' meant the same as 'least like' in accordance with a specific accuracy criterion, this novel word, with its derived meaning, was then inserted into the rule for responding on the subsequent MTS task (i.e., choose the comparison image that is *beda* the sample image).

Participants initially responded on 100 MTS trials in which the rule was reinforced by the task contingencies (i.e., awarded one point each time the participants chose the least like comparison). Immediately following the 100 trials, the contingency reversed and now participants were required to complete another 50 MTS trials in which the rule was punished by the task contingencies (i.e., losing one point for choosing the least like comparison). The main question the researchers sought to address was whether levels of persistent rule-following would differ between participants responding in accordance with the direct rule or a rule that involved a novel derived relation.

Participants who had been provided with the direct rule (in Experiment 2 of the study) persisted for significantly more trials than the derived-rule participants. The authors speculated that this difference may have emerged because in the direct rule group, the entire rule *and* the key phrase "*least like*" involved a history of natural language, the meaning of which would almost certainly have been derived many times across numerous contexts. In the derived rule condition, however, the key phrase participated in a single novel derived relational network (i.e., 'least like' *same as* 'eagsula' *opposite to* 'un' *opposite to* 'beda', therefore 'least like' *same as* 'beda'), and thus the meaning was derived in the individual's history only relatively few times. Within RFT, the extent to which a particular pattern of derived relational responding has been

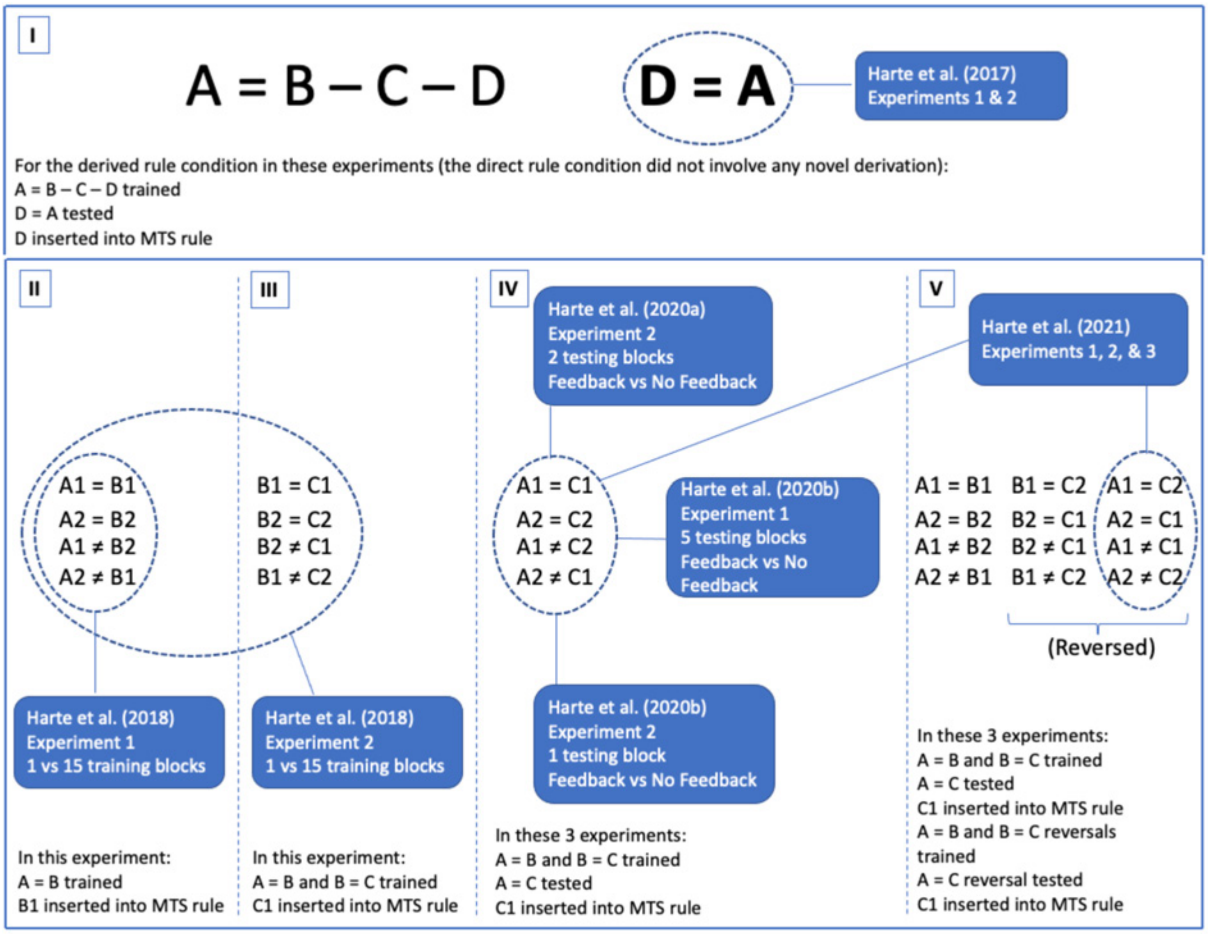

**Figure 2.** A graphical illustration of the relations trained and tested across all Harte and colleagues' studies described currently

**Note**. The '-' in panel one of the figure is used to denote 'opposite to', while the '=' and '≠' are used throughout to denote 'same as' and 'not the same as' respectively. For the ease of the reader, the Harte, Barnes-Holmes, Barnes-Holmes, McEnteggart, et al. (2020) paper has been abbreviated to 'Harte et al. (2020a)', while the Harte, Barnes-Holmes, Barnes-Holmes, McEnteggart (2020) paper has been abbreviated to 'Harte et al. (2020b)'. The reader should also note that the relations within the broken lined circles highlight the specific relations that were the primary focus of the experimental manipulation of any particular experiment. For example, in Experiment 2 of Harte et al. (2018), A=B and B=C relations were trained for either 1 or 15 blocks before the C1 relation ('beda') was inserted into the rule for responding on the MTS task. In Harte et al. (2020a), however, while the A=B and B=C relations were first trained to criteria, it was the derived A=C relations that were the focus of the experimental manipulation (i.e., Feedback versus No Feedback).

emitted previously is considered an important feature of derived relating, referred to as level of derivation (e.g., Barnes-Holmes et al., 2017, 2020). The first time a relation is derived, it is derived directly from an initial relation and is therefore said to be *high in derivation*. Each time a relation is derived, it is thought of as *reducing* in derivation as it acquires its own history, thus making it less and less derived from the original relation (relatively *low in derivation*). The researchers therefore suggested that a rule that involved lower levels of derivation (i.e., direct rule) may produce greater persistent rule-following than a rule that involved higher levels of derivation (i.e., derived rule). On balance, the researchers acknowledged that this interpretation was post-hoc because levels of derivation had not been manipulated directly in the study (i.e., it was assumed that the direct-rule condition involved a relatively low level of derivation). A subsequent study, therefore, aimed to test the interpretation more directly.

**Exploring the impact of derivation**. In a follow-up study, Harte et al. (2018) manipulated derivation by providing different amounts of training to participants (i.e., there was no direct rule condition). In the first of two experiments, a basic network involving a critical mutually entailed relation between the key phrase *least like* and novel word *beda* was trained to different levels (see panel two of Figure 2 in which this critical relation is presented as A1 = B1). That is, participants in one group were required to produce this relational response for 1 block of training trials (high derivation) while another group received 15 blocks of training trials (low derivation). This novel word was then inserted into the rule for responding (i.e., choose the comparison image that is *beda* the sample image) on the same MTS task as described above. A second experiment partially replicated this procedure but the novel network trained now involved establishing a critical combinatorially entailed relation (least like = XXX = beda; illustrated as A1 = B1 = C1 in panels two and three of Figure 2) for one versus 15 blocks. In general, the results indicated that lower levels of derivation (in terms of larger numbers of training trials) produced greater rule persistence than higher levels for both mutually and combinatorially entailed relations.

**Exploring the impact of feedback**. Two subsequent studies explored the impact of feedback for the derived relations on persistent rule-following (Harte, Barnes-Holmes, Barnes-Holmes, McEnteggart et al., 2020; Harte, Barnes-Holmes, Barnes-Holmes, & McEnteggart, 2020; note that these studies are denoted as Harte et al. 2020a and 2020b respectively in Figure 2). In both studies, the critical combinatorially entailed relations involving 'least like' and 'beda' were trained and tested (i.e., train least like = XXX and XXX= beda, test for least like = beda; illustrated as A1, B1 and C1 respectively in panels two, three, and four in Figure 2), before inserting the novel word into the rule for responding on the same MTS task as above. Unlike the previous study by Harte et al. (2018), participants were provided with test trials for the combinatorially entailed relations (see panel four of Figure 2) and feedback was manipulated (i.e., for some participants performance feedback followed each trial but for other participants no differential performance feedback was given). Critically, all participants were required to show similar levels of accurate responding during these tests trials and also on the initial rule-following MTS trials. Thus, any differences that emerged between the feedback and no-feedback conditions could not be attributed to the feedback simply producing better performances during the critical test phases.

Level of derivation was also manipulated across the studies but held constant within each. That is, participants had the same amount of opportunities to derive the novel relation within each experiment, but this varied across experiments. For example, in one experiment participants had a total of five test blocks to derive the critical relation but only one test block in another experiment. Thus, feedback was manipulated *within* each experiment, with level of derivation remaining constant, but *across* each experiment level of derivation varied. Results showed that providing feedback produced increased rule persistence when derivation was high (i.e., fewer test trials), but not when derivation was relatively low (i.e., more test trials). That is, it appeared that the less derived the rule became, the less impact feedback had on persistent rule following. In contrast, when the rule was relatively high in derivation, feedback significantly impacted rule persistence. In other words, it appears that when participants have derived a relation many times, performance feedback for that relation has a limited impact on subsequent rule-following. More informally, the more frequently we derive a relation the less we rely upon external feedback to establish the "truth value" of that relation, at least in terms of persistent rule-following. Indeed, in another recently published study it appeared that it was possible to reduce persistent rule-following for a "strong" relative to a "weak" derived relation by undermining the "truth value" of feedback itself (Bern et al., 2020).

**Exploring relational flexibility and feedback with individual participants**. While all of the research described thus far employed group designs, a recent study sought to explore persistent rule-following as derived relational networks using a single-participant design (Harte et al., 2021). In so doing, the researchers aimed to extend the technical analyses described above while striving for relatively increased levels of behavioral prediction-and-influence, with precision, scope and depth. Specifically, the researchers explored the extent to which flexibility in reversing derived relations would impact upon MTS persistent rule-following.

The first of three experiments focused entirely on assessing participants' ability to reverse derived relations given previous findings that had suggested that reversing such relations may be difficult (e.g., Pilgrim & Galizio, 1990, 1995). Specifically, three participants were first trained and tested on a relational network comprised of two combinatorially entailed relations (i.e., train 'least like'(A1) = 'XXX'(B1) = 'beda'(C1); and 'most like'(A2) = ']][['(B2) = 'sarua'(C2), test 'least like' = 'beda' and 'most like' = 'sarua'; illustrated across panels two, three, and four of Figure 2). Next, two of the mutually entailed relations were reversed and the new combinatorially entailed relations were tested (i.e., train 'least like' = 'XXX' = '*sarua*' and 'most like' = ']][[' = '*beda*', test 'least like' = 'sarua' and 'most like' = 'beda'; see panel five of Figure 2). The objective here was to determine if reversing mutually entailed relations produced reliable (i,e, flexible) reversals in combinatorial relations. The experiment involved three such reversals to determine if a relatively high level of flexibility could be achieved using these procedures. Results showed that all three participants successfully produced a test performance that was in accordance with the most recently trained relations, thus demonstrating clear flexibility in derived relating.

In the second experiment, three additional participants were first trained and tested on the same derived relational network as Experiment 1. Immediately afterward, participants were provided with a rule for responding on a subsequent MTS task that involved a part of the trained and tested network (i.e., choose the comparison image that is *beda* to the sample image; note that all participants had successfully derived that 'beda' meant 'least like'). MTS feedback contingencies thus reinforced responding in accordance with this rule. Participants were then trained and tested on the reversed network (as in Experiment 1), after which they were given the same rule for MTS responding. Thus, the reversed derived relations now entailed that 'beda' meant 'most like'; crucially the MTS feedback contingencies did not reverse and thus following the rule was now punished. The main aim was to assess the extent to which reversed derived relations would override the MTS task feedback contingencies. Results showed that, despite successfully reversing the derived relations, as was observed in Experiment 1, all three participants tended to respond readily in accordance with the MTS task feedback contingencies (i.e., they appeared to "ignore" the rule as specified by the reversed derived relations).

The third and final experiment partially replicated Experiment 2, but the MTS task also involved a reversal of feedback contingencies. Specifically, upon completing the same training and testing of the reversed network as in Experiments 1 and 2, the first half of the MTS trials were also reversed. That is, the MTS trials, initially, were consistent with the reversed derived rule but then reversed (for the final half of the trials), thus punishing responding in accordance with the derived rule. Results showed that all three participants showed evidence of following the derived rule even after the MTS contingencies reversed (i.e., participants no longer simply ignored the rule).

The authors offered the following interpretation of their findings. Specifically, two competing relational networks may have been at play in Experiments 2 and 3: the rule from the derived training and testing procedure, and a rule that emerged through direct contact with the MTS task contingencies. In Experiment 2, the MTS contingencies remained unchanged throughout the experiment, and therefore the MTS rule/network provided a relatively strong (i.e., reliable) source of behavioral control. In contrast, the network resulting from the derived relations was reversed multiple times throughout the experiment, thus participants responded in accordance with the network that never changed (i.e., the one generated by the MTS feedback contingencies). As a result, participants failed to show any evidence of derived rule-persistence in the face of competing MTS contingencies. In Experiment 3, however, the MTS contingencies also reversed within the experiment, and thus the reliability of the network generated by those contingencies was reduced relative to Experiment 2. As a result, the relative controlling properties of the derived rule may have been greater in Experiment 3, which resulted in evidence of rule-persistence for the derived network/rule. Once again, the authors acknowledged that the post-hoc explanation was quite speculative, but it nonetheless illustrated the complexities and subtleties likely involved in the study of (persistent) rule-following, when rules are interpreted as (derived) relational networks.

## Concluding comments

Skinner (1966) first proposed the concept of rule-governed behavior in his chapter on an operant analysis of problem solving. This conceptual development generated a relatively rich behavior-analytic literature on rule-governed behavior, part of which focused on the extent to which rule-following may generate insensitivity to "direct" contingencies of reinforcement. In 1971, Sidman identified an effect that he and his colleagues later labelled stimulus equivalence, and this effect was used to produce a functional-analytic interpretation of how rules specify reinforcement contingencies (see Sidman 1994). Only recently, however, have researchers attempted to bring together the study of rule-governed behavior and stimulus equivalence, and derived relations more generally, to help us better understand how rules may indeed undermine sensitivity to direct contingencies of reinforcement (see Harte, Barnes-Holmes, Barnes-Holmes, & Kissi, 2020, for an extended discussion).

The relatively recent program of research described above that has emerged in this vein highlights what appear to be complex and subtle effects in exploring the impact of derivation, feedback, and relational flexibility on persistent rule-following, at both the group and individual participant level. The findings demonstrate the fruits of considering rules as derived relational networks and that the behavior-controlling properties of these networks in the context of persistent rule-following are subject to the influence of variables such as derivation, feedback, and flexibility. For example, the findings of Harte et al. (2017, 2018) indicated that the more often a relation involved within a network had been derived previously, the more likely the network would control responding in the face of competing reinforcement contingencies. In addition, subsequent studies by Harte and colleagues indicated that the provision of corrective feedback for deriving a relation, involved within a network or rule, tends to produce greater levels of rule-persistence than providing no feedback (Harte, Barnes-Holmes, Barnes-Holmes, McEnteggart et al., 2020). On balance, evidence also suggests that the more often the relation is derived the less impact impact corrective feedback appears to have on rule-persistence (Harte, Barnes-Holmes, Barnes-Holmes & McEnteggart, 2020). Furthermore, recent research that has extended this work in the context of individual participants has demonstrated that inducing relational flexibility in deriving a network appears to undermine the extent to which that network subsequently controls rule-following that involves using the network in a rule (Harte et al., 2021). However, this effect is itself undermined if flexibility is induced in the task that is used to test for rule persistence itself. In other words, if the feedback contingencies for both deriving part of a rule, and for a task that tests rule-following itself, are reversed, then the impact of the performance feedback itself appears to lose its behavior-controlling properties.

The implications of the foregoing analyses, and in particular the interpretation of findings provided by Harte et al. (2021), suggest that a shift in perspective may be required in exploring rule-persistence and so-called contingency insensitivity. In effect, these findings suggest that it may not necessarily be the case that rules undermine sensitivity to direct contingencies of reinforcement, but rather that a derived network (rule 1) competes with another network (rule 2), which was generated based on direct contact with the feedback contingencies. This suggestion may be important going forward for how we as researchers proceed in conducting basic and applied experimental analyses of so-called persistent rule-following or contingency insensitivity. In other words, it may not be that rules undermine sensitivity per se, but rather that multiple relational networks (or rules) compete with each other in terms of controlling behavior in specific contexts or tasks. In this sense, it may be useful to conceptualize instances of persistent rule-following as involving the relating of relational networks, in which the relative coherence (or truth value) of one network is weighed against the coherence (or truth value) of another network. Viewing persistent rule-following as involving competing relational networks then begs questions concerning to what extent does one network acquire functionally distinct behavior controlling properties over another network? And as an extension, what are the variables important in predicting-and-influencing this control?

In adopting the foregoing conceptual analysis it could be argued that it remains very close to a Skinnerian view that insensitivity to reinforcement contingencies is not a demonstration of lack of operant control, but rather evidence of competition between two separate classes of generalized operant behavior (i.e., a relational network that involved a specific trained and tested derived relation versus a relational network that was generated from direct interaction with the feedback contingencies for a specific task). In any case, if nothing else, the foregoing arguments serve to highlight the highly complex relational phenomena that appear to be involved in rule-governed behavior. And as such, a great deal of experimental research is needed if RFT is to contribute towards an operant understanding of the extent to which rules (as relational networks) control behavior in any given instance. Indeed, this work seems particularly important if we are to sufficiently advance our understanding of how excessive or inflexible rule-following plays a key role in human psychological distress, as has been long argued by the ACT literature.

## References

Albuquerque, L. C., Reis, A. A., & Paracampo, C. C. P. (2006). Efeitos de uma história de reforço contínuo sobre o seguimento de regras. *Acta Comportamentalia*, *14*, 47-75.

Albuquerque, L. C., Souza, D. G., Matos, M. A., & Paracampo, C. C. P.(2003). Análise dos efeitos de histórias experimentais sobre o seguimento subsequente de regras. *Acta Comportamentalia*, *11*, 87-126.

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

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. 103-118). Plenum.

Bern, R., Persdotter, T., Harte, C., & Barnes-Holmes, D. (2020). Relational coherence and persistent rule-following: The impact of targeting coherence in a 'non-critical' component of a relational network. *The Psychological Record.* Advance online publication. https://doi.org/10.1007/s40732-020-00414-1

Catania, C. (1979). *Learning.* Englewood Cliffs: Prentice Hall.

Catania, C. (1999) *Aprendizagem.* Artmed: Porto Alegre.

Cortez, M. C. D., & Reis, M. J. D. (2008). Efeitos do controle por regras ou pelas contingências na sensibilidade comportamental. *Revista Brasileira de Terapia Comportamental e Cognitiva*, *10*(2), 143-155.

de Almeida, J. H., Cortez, M. D., & de Rose, J. C. (2020) The Effects of Monitoring on Children's Rule-Following in a Computerized Procedure. *The Analysis of Verbal Behavior*, *37,* 1-1. https://doi.org/10.1007/s40616-020-00130-5

de Rose, J. C., & Almeida, J.H. (2019). Aprendendo mais do que foi ensinado. In P. G. Soares, J. H. Almeida & C. R. X. Cançado (Eds.), *Experimentos clássicos em análise do comportamento*, (pp. 211-225). Instituto Walden4.

de Rose, J. C. (2012). Classes de estímulos: implicações para uma análise comportamental da cognição. *Psicologia: Teoria e Pesquisa, 9*(2), 283-303.

de Rose, J. C., de Souza, D. G. Rossito, A. L., & de Rose, T. M. (2012). Aquisição de leitura após história de fracasso escolar: equivalência de estímulos e generalização. *Psicologia: Teoria e Pesquisa, 5*(3), 325-346.

de Rose, J. C., Gil, M. S. C. A., & de Souza, D. D. G. (2014). *Comportamento simbólico: bases conceituais e empíricas*. São Paulo, SP: Cultura Acadêmica (Unesp).

Dougher, M. J., Hamilton, D. A., Fink, B. C., & Harrington, J. (2007). Transformation of the discriminative and eliciting functions of generalized relational stimuli. J*ournal of the Experimental Analysis of Behavior*, *88*, 179-197. https://doi.org/10.1901/jeab.2007.45-05

Dymond, S. & Rehfeldt, R. A. (2000). Understanding complex behavior: The transformation of stimulus functions. *The Behavior Analyst, 23*, 239-254. https://doi.org/10.1007/BF03392013

Friman, P. C., Hayes, S. C., & Wilson, K. G. (1998). Why behavior analysts should study emotion: The example of anxiety. *Journal of Applied Behavior Analysis*, *31*(1), 137-156. https://doi.org/10.1901/jaba.1998.31-137

Harte, C. & Barnes-Holmes, D. (2021). 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, 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. *Behavioural Processes, 157,* 36-46. https://doi.org/10.1016/j.beproc.2018.08.005

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., Barnes-Holmes, Y., & McEnteggart, C. (2020). Exploring the impact of coherence (through the presence versus absence of feedback) and levels of derivation on persistent rule-following. *Learning and Behavior*. Advance online publication. https://doi.org/10.3758/s13420-020-00438-1

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. (1989). *Rule-governed behavior: Cognition, contingencies, and instructional control*. Plenum.

Hayes, S. C. (1991). A relational control theory of stimulus equivalence. In L. J. Hayes & P. N. Chase (Eds.), *Dialogues on Verbal Behavior* (pp. 19–40). Reno, NV: Context Press.

Hayes, S.C. (2004). Acceptance and commitment therapy, relational frame theory, and the third wave of behavioral and cognitive therapies. *Behavior Therapy, 35*, 639–665. https://doi.org/10.1016/S0005-7894(04)80013-3

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., Strosahl, K., & Wilson, K.G. (1999). *Acceptance and Commitment Therapy: An Experiential Approach to Behavior Change.* Guilford Press.

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

Kohlenberg, R. J., & Tsai, M. (1991). *Functional analytic psychotherapy*. Springer.

Leonardi, Jan Luiz. (2015). O lugar da terapia analítico-comportamental no cenário internacional das terapias comportamentais: um panorama histórico. *Perspectivas em análise do comportamento*, 6(2), 119-131. https://dx.doi.org/10.18761/pac.2015.027

Linehan, M. M. (1993). *Cognitive-behavioral treatment of borderline personality disorder.* Guilford.

Lucena-Santos, P.; Pinto-Gouveia, J. & Oliveira, M.S. (2015). Primeira, segunda e terceira geração de terapias comportamentais. In: P. L. Santos, J. P. Gouveia, M. S. Oliveira (Orgs.) *Terapias comportamentais de terceira geração.* Novo Hamburgo: Sinopsys, pp. 29-58.

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. http://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 Behaviour, 102* (1), 66-85. http://doi.org/10.1002/jeab.95.

Paracampo, C. C. P., & de Albuquerque, L. C. (2005). Comportamento controlado por regras: revisão crítica de proposições conceituais e resultados experimentais. *Interação em psicologia,* 9(2), 227 - 237. http://dx.doi.org/10.5380/psi.v9i2.4798

Paracampo, C.C.P., Albuquerque, L.C., Farias, A.F., Carvalló, B.N., & Pinto, A.R. (2007). Efeitos de consequências programadas sobre o comportamento de seguir regras. *Interação em Psicologia*, 11, 161-173. http://dx.doi.org/10.5380/psi.v11i2.7850

Paracampo, C. C. P., Albuquerque, L. C., Mescouto, W. A. A., & Farias, A. F. Efeitos de perguntas e de respostas às perguntas sobre o seguir regras apresentadas em uma história infantil. Psicologia: Teoria e Pesquisa, 29, 369-379. https://doi.org/10.1590/S0102-37722013000400003

Perez, W. F., Nico, Y. C., Kovac, R., Fidalgo, A. P., & Leonardi, J. L. (2013). Introdução à teoria das molduras relacionais (Relational Frame Theory): principais conceitos, achados experimentais e possibilidades de aplicação. *Perspectivas em Análise do Comportamento,* 4 (1), 32-50. https://doi.org/10.18761/perspectivas.v4i1.105

Pilgrim, C. & Galizio, M. (1995). Reversal of baseline relations and stimulus equivalence: I. Adults. *Journal of the Experimental Analysis of Behavior, 63*, 225-238. https://doi.org/10.1901/jeab.1995.63-225

Pilgrim, C. & Galizio, M. (1990). Relations between baseline contingencies and equivalence probe performances. *Journal of the Experimental Analysis of Behavior, 54*, 213-224. https://doi.org/10.1901/jeab.1990.54-213

Reis, M. J. D., Perez, W. F., & Arantes, A. K. L. (2010). Perdas e ganhos: Efeitos de consequências programadas sobre o seguimento de regras. *Revista Brasileira de Análise do Comportamento*, 6, 125-148. http://dx.doi.org/10.18542/rebac.v6i2.1115

Santos, J. G. W., Paracampo, C. C. P., & Albuquerque, L. C. (2004). Análise dos efeitos de histórias de variação comportamental sobre o seguimento de regras. *Psicologia: Reflexão e Crítica*, 17 (3), 413-425. http://dx.doi.org/10.1590/S0102-79722004000300013

Segal, Z. V., Williams, J. M. G., & Teasdale, J. D. (2002). *Mindfulness-based cognitive therapy for depression: A new approach to preventing relapse.* Guilford.

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

Sidman, M. (1994). *Equivalence Relations and Behaviour: A Research Story*. Boston, MA: Authors Cooperative.

Skinner, B.F. (1957). *Verbal behavior*. New York: Appleton-Century-Crofts.

Skinner, B.F. (1966). An operant analysis of problem solving. In B. Kleinmuntz (Ed.), *Problem-Solving: Research, Method, and Theory*, (pp. 225-257). Wiley.

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

Torneke, N. (2010) *Learning RFT: An Introduction to Relational Frame Theory and Its Clinical Applications*. Oakland: New Harbinger Publications.

Torneke, N., Luciano, C., Valdivia-Salas, S. (2008). Rule-governed behavior and psychological problems. *International Journal of Psychology and Psychological Therapy*, *8* (2), 141 – 156.

Zettle, R.D. & Hayes, S.C. (1982). Rule-governed behavior: A potential theoretical framework for cognitive behavior therapy. In P.C. Kendall (Ed.), *Advances in Cognitive Behavioral Research and Therapy* (pp. 73-118). Academic.

**Informações do Artigo**

**Histórico do artigo:**
Submetido em: 20/01/2021
Primeira decisão editorial: 07/04/2021
Aceito em 05/05/2021
Editor: William F. Perez