BEHAV ANALYST (2017) 40:65–74
DOI 10.1007/s40614-017-0099-4

COMMENTARY

RFT as a Functional Analytic Approach to Understanding the Complexities of Human Behavior: A Reply to Killeen and Jacobs

Ian Stewart[1]

Published online: 9 May 2017

Killeen and Jacobs (2017) argue that Skinner's three-term contingency by itself fails to deal adequately with complexities of human motivation including such phenomena as purpose, wants, interests, attitudes, and beliefs. Their proposed solution is an expansion of this concept to include the organism as a moderator of the environment-behavior relationship. I would argue that they are correct in their suggestion that operant learning as traditionally known is inadequate as an account of complex human behavior but I disagree with their proposed solution, because it represents a departure from Skinner's account at a fundamental philosophical level that I believe would be destructive and unnecessary. In this reply to Killeen and Jacobs, I first outline my opposition to the suggestion that the organism be imported into the three-term contingency. I then proceed to describe a more philosophically conservative yet empirically promising approach to understanding and dealing with the complexities of human behavior.

The truly distinctive character of behavior analysis arguably lies in its philosophically pragmatic roots. The goal of this approach is not uncovering objective truth but achieving prediction and influence over behavior, and the key conceptual tool used to achieve this end is the operant, defined in terms of relations between behavior and its current and historical context. This definition facilitates a pragmatic orientation by allowing a focus on manipulable elements of this environment. The scientist or practitioner cannot manipulate behavior itself directly but she can manipulate the context (including both antecedents and consequences) in which the behavior takes place and thus she can bring about behavioral change.

The idea around which the Killeen and Jacobs article is based, that reinforcement is not an inherent property of any particular stimulus object accords with the pragmatic orientation described above and in particular with the definition of the operant in terms of behavior-environment relations. Given the latter, both behavior (responses) and

✉ Ian Stewart
ian.stewart@nuigalway.ie

1 National University of Ireland, Galway, Ireland

environment (stimuli) are defined functionally, each with respect to the other, rather than in formal terms. As a result, for example, neither food nor any other category of object would be characterised as a reinforcer in and of itself; instead, a particular stimulus object such as food may function as a reinforcer in particular historical and current contextual conditions and to the extent that we can identify manipulable aspects of those conditions then we may achieve greater or lesser influence over the behavior involved. The fact that functions such as reinforcement are fundamentally determined by context in this way is not a weakness of this approach but is instead a defining feature of behavior analysis on the basis of which it has been extremely successful in a number of applied domains.

Killeen and Jacobs (2017) argue that we need to add the organism into the operant as an additional element so as to facilitate prediction and influence over behavior and in particular the complexities of human behavior. However, the organism is simply part of the locus at which the effects of historical and current contingencies may be recorded and it is these contingencies and their behavioral effects that are technically important and on which we as behavior analysts need to focus, not any particular spatiotemporal locus wherein the effects occur. Giving the organism special status as something different from and additional to either act or context would be a step away from the pragmatic character of behavior analysis as described above and towards a more mechanistic, essentialist conception. This is because it would introduce the organism as an entity requiring definition in formal rather than functional terms and one which would likely depend on biological rather than behavioral science, and hence would represent a reductionist move (apart from this, there is in fact at the present time no agreed technical definition of the term *organism* even within biological science so this would introduce an additional layer of conceptual problems; see Pepper & Herron, 2008; Roche & Barnes, 1997). Furthermore, given the fact that description of the formal structure of the organism as a central hub of influence over behavior might easily come to dominate theorising and empirical research, it seems quite possible that behavior analysis thus changed might become indistinguishable from cognitive psychology.

Hence, designating the organism as a formal element of the operant would be a radical and, in my opinion, damaging move for behavior analysis. At the same time, I recognise that Killeen and Jacobs' (2017) suggestion is an attempt to help the field step up to the challenges involved in obtaining prediction and influence in particular with respect to complex human behavior. From my perspective, there is no need to take the step suggested by these authors in order to do this, however. A much more promising route in this direction requires, not requires abandoning the pragmatic foundations of operant psychology, but instead following the data gleaned from empirical research on a particular type of operant, namely, derived relational responding. As it turns out, these data do lead to a relatively radical conclusion. Despite this, however, this approach remains firmly rooted in both pragmatism and operant psychology.

Derived relational responding is an empirical phenomenon in which training particular relations between arbitrary stimuli leads to the derivation of additional untrained relations between those stimuli in accordance with a particular overarching pattern. The most widely researched example is derived equivalence (e.g., Sidman, 1971; Sidman et al., 1982; Sidman & Tailby, 1982) wherein participants taught certain unidirectional relations among physically dissimilar stimuli using a match-to-sample protocol derive further relations in accordance with an overarching pattern of derived sameness. A number of other patterns of derived relational responding have also been studied including, for example, derived opposition (e.g., if trained that A is opposite B then participants derive that B is opposite A; Steele & Hayes, 1991; Stewart et al., 2015), comparison (e.g., if trained that A is more than B, participants derive that B is less than A; Dymond & Barnes, 1995; Berens & Hayes, 2007), distinction (e.g., Roche & Barnes, 1996), hierarchy (e.g., Gil, Luciano, Ruiz, & Salas, 2012; Slattery & Stewart, 2014), temporality (O'Hora, Pelaez, Barnes-Holmes, & Amesty, 2005; O'Hora et al., 2008), deixis (e.g., McHugh, Barnes-Holmes, & Barnes-Holmes, 2004; Rehfeldt, Dillen, Ziomek, & Kowalchuk, 2007; Weil, Hayes, & Capurro, 2011), and analogy (e.g., Barnes, Hegarty, & Smeets, 1997; Stewart, Barnes-Holmes, Roche, & Smeets, 2002; for comprehensive reviews, see Dymond & Roche, 2013; Stewart, 2016; Zettle, Hayes, Barnes-Holmes, & Biglan, 2016).

Derived relational responding has received extensive research attention for a number of reasons, including both that it is difficult to account for in terms of traditional learning principles as well as that it seems closely related to human language (see e.g., Hayes, Barnes-Holmes, & Roche, 2001a; Hayes et al., 2001b; Stewart & Roche, 2013). With respect to the latter, one particularly important effect associated with derived relational responding that seems relevant as regards linguistic meaning is *transformation of function* whereby the psychological functions of a stimulus change by virtue of being in a derived relation with one or more other stimuli. For example, Dougher, Augustson, Markham, Greenway, and Wulfert (1994) demonstrated one of the first examples of transformation of functions when they showed the transfer of aversive functions to a previously neutral stimulus based on derived equivalence with a directly shock-conditioned stimulus. In a more recent study, Dougher, Hamilton, Fink, and Harrington (2007) showed transformation of functions through derived comparison relations. They first induced participants to derive a relation of "C more than B" in two arbitrary stimuli before pairing B with shock; subsequently, most participants showed higher levels of physiological responding to C than to B, despite the fact that C had not been paired with shock. These empirical effects are examples of the generativity of the phenomenon of derived relational responding as well as the close similarity between this phenomenon and linguistic meaning.

One theory that has been forwarded to explain the origins of derived relational responding and its correlation with human language is relational frame theory (RFT; Hayes et al., 2001a, 2001b; Dymond & Roche, 2013). This approach explains both derived relational responding and language as instances of an operant repertoire referred to as arbitrarily applicable relational responding (AARR). To explain, many species can show what is referred to as non-arbitrary relational responding in which they relate stimuli based on their physical properties (e.g., learning to consistently respond to a comparison that is physically larger or physically smaller than the sample). However, typical members of the human species exposed to contingencies provided by their socioverbal community can learn in addition to show arbitrarily applicable relational responding. This is primarily based not on non-arbitrary or formal relations between the stimuli related but on aspects of the context (referred to within RFT as "contextual cues") that specify the relation such that the relational response can be brought to bear on any relata regardless of their non-arbitrary properties (see Stewart & McElwee, 2009).

As an example of how derived relational responding can be explained as AARR, consider derived comparative relations. Imagine I teach a verbally able child that, "Mr. X is taller than Mr. Y and Mr. Y is taller than Mr. Z" and imagine that the child is then able to derive multiple new relations including, for instance, "Mr. X is taller than Mr. Z" and "Mr. Z is shorter than Mr. X." The child is able to do this despite the fact that the latter performances have not been explicitly taught and that the non-arbitrary properties of the stimuli involved do not support these performances (for instance, it is not obvious that the stimulus "Mr. X" is physically taller than the stimulus "Mr. Z"). From an RFT perspective, what is happening is that I am presenting the child with a contextual cue (i.e., "taller") that their learning history has previously established as controlling a particular pattern of generalised relational responding. In this history, the child probably starts by learning contextually controlled non-arbitrary relational responding, involving choosing the physically taller of two objects in the presence of "taller" and the physically shorter in the presence of "shorter." Then, following multiple exemplars of the reinforcement of this pattern in the presence of these cues, the relational response becomes abstracted such that it can be applied in conditions in which there is no obvious formal relation. At that point, the pattern can be brought to bear on any arbitrarily chosen set of stimuli no matter what their non-arbitrary properties or the non-arbitrary relations between them such that all of the stimuli are brought into a coherent set of relations with each other. In this particular example, it is brought to bear on the arbitrary stimuli "Mr. X," "Mr. Y," and "Mr. Z".

Arbitrarily applicable relational responding is also referred to as relational framing (hence *relational frame* theory), based on the metaphor of a picture frame that can contain varying content, and as this term is less of a mouthful, and perhaps more intuitively understood, this term will be used from here on. Relational framing comes in multiple forms, supporting the multiple patterns of derived relations listed earlier, and indeed, many studies focused on equivalence and most that have examined non-equivalence type derived relational responding have been RFT-inspired and have thus been explicitly interpreted as examples of the demonstration of relational framing as well as of derived relations.

Despite this variation, all forms of relational framing share a number of key properties. The first, mutual entailment, refers to the finding that a relation trained in one direction entails derivation of a second relation in the other (e.g., if $A > B$, then $B < A$). The second, combinatorial entailment, refers to the fact that taught relations are combined to derive new relations (e.g., if $A > B$ and $B > C$, then $A > C$ and $C < A$). Third is the so-called transformation of functions effect referred to earlier in which a stimulus in an arbitrarily applicable relation with another stimulus can change its functions simply through being in that relation (e.g., if $A > B > C$ and if B is conditioned as aversive, then A will become less aversive than B and C more aversive; see Dougher et al., 2007).

These symbolic and generative properties strongly suggest that relational framing constitutes the key operant underlying human language and cognition (e.g., Hayes & Hayes, 1992). Furthermore, considerable work has supported this argument by (a) showing that relational framing is not just correlated with language but emerges before development of an advanced linguistic repertoire (e.g., Luciano, Gómez-Becerra, & Rodríguez-

Valverde, 2007), (b) by demonstrating that training relational framing with abstract stimuli can substantially boost verbal and intellectual ability (e.g., Cassidy, Roche, Colbert, Stewart, & Grey, 2016; Hayes & Stewart, 2016), (c) by reporting controlled empirical demonstrations of the origins and development of various forms of relational framing (e.g., Berens & Hayes, 2007), (d) by convincingly showing that relational framing meets criteria as an operant behavior (e.g., Healy, Barnes-Holmes, & Smeets, 2000), and (e) by demonstrating relational framing as a useful functional analytic unit for the analysis of human language and cognition (Dymond & Roche, 2013).

From the RFT perspective, then, relational framing is the process that underlies language (see e.g., Hayes et al., 2001a, 2001b). As befits such a process, relational framing is potentially extremely generative. Any stimulus can be related to any other stimulus in accordance with any relational frame, including those listed above as well as others, and have its functions transformed accordingly. Of course, the fact that relational framing is arbitrarily applicable does not mean that it is arbitrarily applied, and indeed natural language has its greatest use in the description and analysis of the formal features of our environment. Nevertheless, the empirically established and analyzable process of relational framing can explain the generativity and flexibility that language facilitates, from trivial everyday examples (e.g., naming a pet dog) to the extremes of artistic creativity (e.g., writing a best-selling novel) or scientific innovation (e.g., developing a comprehensive empirically based theory of the origins of the universe).

From this perspective, any objects or events that are relationally framed become verbal for us—they become part of the world as known through relational frames. As we frame objects, events, and people through our interactions with the socioverbal community, we elaborate our network of related stimuli, and through transformation of functions, the world increasingly takes on new verbally derived functions. The expansion of this network starts when we first learn to frame words and objects as the same and probably continues throughout most of our lives. The well-documented "language explosion" between the ages of 2 and 3 is an obvious and salient example of the elaboration of the relational network. This typically occurs around the time that children have likely acquired the ability to frame in accordance with a few simple relations, allowing them to derive multiple novel relations among an expanding set of named objects and events. As children grow into adulthood, continued verbal interactions produce an increasingly complex and multirelational network involving vast numbers of different objects and events and the relations between them. For human beings, everything we encounter and think about, including ourselves, our thoughts and emotions, our prospects, other people, and our environment, becomes part of this elaborate verbal relational network. Thus, for human beings, the world is verbal, and we can never escape from language except under very unusual circumstances.

With respect to the focus of the Killeen and Jacobs (2017) article, two important points can be made vis a vis this RFT analysis. First, it affords a bottom-up behavior analytic approach to explaining complex human behavior including the phenomena cited as important by Killeen and Jacobs. Indeed, there are currently over 200 journal articles and several books describing RFT theory and empirical research with respect to a variety of key domains of human psychology including language development, linguistic generativity, rule-following, analogical reasoning, intelligence, theory of mind, psychopathology, attitudes/beliefs, and motivation (e.g., Dymond & Roche, 2013; Hayes et al., 2001a, 2001b; Hayes & Stewart, 2016; Stewart, 2016). I will say a little more on the latter two areas, which are ones explicitly mentioned by Killeen and Jacobs.

Regarding attitudes and beliefs, the decade just passed has seen a very substantial quantity of RFT work in this area. Over the last 20 years, many psychology researchers have focused on measuring so-called implicit attitudes with the rationale that doing so might facilitate improved prediction of behavior compared to (explicit) self-report in a range of important domains (see Nosek, Hawkins, & Frazier, 2011). A majority of this research has been carried out by researchers operating within a mentalistic paradigm, who have largely focused on the specific mental constructs that are assumed to mediate between implicit attitudinal behavior and the environment. Furthermore, the predominant conceptualization of implicit attitudes up until recently from within this orientation was that they are based on mental associations instantiated somehow in the mind/brain. Within the last decade, however, RFT researchers have offered a novel functional analytic conceptualization of implicit attitudes in terms of relational framing (see Barnes-Holmes, Barnes-Holmes, Stewart, & Boles, 2010; Hughes, Barnes-Holmes, & Vahey, 2012). This approach argues (a) that both explicit (i.e., self-report) and implicit attitudes represent instances of the same functionally defined process of relational framing and (b) that differences between them are based on differences with respect to framing complexity (how many stimuli and/or relations are involved in a particular instance of framing) and level of derivation (the more a particular relation has been derived previously, the lower the level of derivation currently involved).

This RFT account presents several potential advantages. First, it is parsimonious in explaining implicit and explicit attitudes as the same basic process differing only with respect to certain qualities (namely complexity and derivation); more specifically, implicit responding is characterised by lower levels of both complexity and derivation than explicit. Second, by arguing for implicit attitudes as relational rather than associational, it allows for a richer conceptualization of this process with greater explanatory power and scope for experimental investigation. The idea of implicit attitudes as relational has in addition foreshadowed recent findings from within the cognitive approach also[1] (see De Houwer, 2014). Third, based on the latter and the functional analytic methods already developed to examine relational responding within RFT, this approach has given rise to a potentially useful new procedure for examining implicit attitudes, namely, the Implicit Relational Assessment Procedure (IRAP; Barnes-Holmes et al., 2006). The IRAP facilitates the exploration of relational framing under time pressure, a key feature of the conditions under which implicit attitudes are typically measured, and has allowed relatively wide-ranging exploration of this domain within a short space of time (see Hughes & Barnes-Holmes, 2013), particularly in the domain of psychopathology (Vahey, Nicholson, & Barnes-Holmes, 2015).

The area of motivation is also one in which RFT/derived relations research has contributed. Such work seems particularly relevant in the current context, given the centrality in Killeen and Jacob's paper of the topic of reinforcement and in particular how we might account for the complexities of human motivation. Several theoretical and empirical papers under the RFT rubric have been produced on this topic (see e.g., Ju & Hayes, 2008; Hayes, Kohlenberg, & Hayes, 1991; Hughes & Barnes-Holmes,

[1] Namely, findings that support the suggestion that so-called "implicit evaluation" effects could be based on rules or propositions (relational networks) rather than simple associations.

2011; Plumb, Stewart, Dahl, & Lundgren, 2009; Whelan & Barnes-Holmes, 2004; Whelan, Barnes-Holmes, & Dymond, 2006). An initial RFT conceptualization of valuing was provided by Barnes-Holmes et al. (2001) and the same topic was covered in greater detail in a more recent paper by Plumb et al. (2009). Both of these pieces describe valuing in terms of transformation of reinforcing and/or establishing functions via relational framing, processes which have also received attention across several empirical articles (see Ju & Hayes, 2008; Hayes et al., 1991; Whelan & Barnes-Holmes, 2004; Whelan et al., 2006). Both Barnes-Holmes et al. (2001) and Plumb et al. (2009) also expanded on their core description of valuing via references to relational networks and hierarchical relational framing, and these phenomena have also themselves received separate empirical treatment in their own right (see O'Hora, Barnes-Holmes, Roche, & Smeets, 2004; O'Hora, Barnes-Holmes, & Stewart, 2014; Gil et al., 2012; Slattery & Stewart, 2014).

The core empirical work in this area has focused on demonstrating how the reinforcing effectiveness of stimuli may be changed via derived relational processes, without the need for direct classical or operant conditioning. RFT studies on transformation of conditioned reinforcing, establishing and evaluative functions respectively provide empirical demonstrations of how derived relational responding can affect motivation with respect to environmental stimuli in potentially wide-ranging and subtle ways. For example, Whelan et al. (2006), who focused on transformation of conditioned reinforcing functions, showed that several arbitrary stimuli acquired reinforcing functions in the absence of direct training by virtue of being in derived comparative (more/less) relations with a single directly conditioned stimulus; furthermore, the relative reinforcing value of these stimuli was correlated with their position in a derived comparative relational network. Ju and Hayes (2008) showed transformation of establishing stimulation functions in both adults and children by showing that formerly neutral stimuli in derived equivalence relations with reinforcing consequences could increase operant responding that produced those consequences.

The RFT analyses and empirical studies described over the last few paragraphs represent a very small proportion of what is a vibrant field of research characterised by novel conceptualizations and methodologies and producing new insights into various important domains of complex human psychology. They underline the potential of RFT as a functional analytic approach to examining the kinds of phenomena that Killeen and Jacobs (2017) mention in their article, and indicate some of the progress made thus far in doing so. As such, they undermine the argument that a new approach which breaks with the fundamental assumptions underlying Skinnerian behavior analysis is needed in order to tackle such phenomena.

While making this point, however, a second point regarding the RFT analysis should also be noted. As indicated earlier, relational framing is indeed an operant in the traditional Skinnerian sense to the extent that no new principles are needed to account for how it is acquired, and thus, in contrast to the Killeen and Jacobs (2017) proposed revision, this approach is firmly rooted in a traditional behavior analytic conceptualization. At the same time, however, as has been noted elsewhere (e.g., Hayes et al., 2001a, 2001b; Stewart, McElwee, & Ming, 2013), while established in the same way as other operants, the outcome of relational framing is utterly unique in that relational framing is a learned process that can affect other behavioral processes. The property of relational framing that confers this capacity is of course transformation of function wherein stimuli in derived relations can acquire novel functions or have previously established functions modified without direct training. This, together with the rapidity and ubiquity of relational framing once it is acquired, together with an environment (i.e., the socioverbal community) in which such responding is strongly supported, means that learning in verbal humans becomes exponentially more diverse and complex than that seen in non-verbal organisms. As such, learning in verbal humans might justifiably be seen as qualitatively distinct from that in non-humans. In this respect, then I agree with Killeen and Jacobs that a different conception is needed with respect to understanding human learning. However, as in the case of RFT, this different conception can remain rooted in behavior analysis and need not jettison the fundamental philosophical assumptions underlying this approach.

Compliance with Ethical Standards

Funding This article was not supported by funding.

Conflict of Interest The author declares that he has no conflict of interest.

References

Barnes, D., Hegarty, N., & Smeets, P. M. (1997). Relating equivalence relations to equivalence relations: a relational framing model of complex human functioning. *The Analysis of Verbal Behavior, 14*, 1–27.

Barnes-Holmes, D., Barnes-Holmes, Y., Haydn, E., Milne, R., Power, P., & Stewart, I. (2006). Do you know what you really believe? Developing the implicit relational evaluation procedure (IRAP) as a direct measure of implicit beliefs. *The Irish Psychologist, 32*(7), 169–177.

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. *Psychological Record, 60*, 527–542.

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*. New York: Plenum Press.

Berens, N. M., & Hayes, S. C. (2007). Arbitrarily applicable comparative relations: experimental evidence for relational operants. *Journal of Applied Behavior Analysis, 40*(1), 45–71.

Cassidy, S., Roche, B., Colbert, D., Stewart, I., & Grey, I. M. (2016). A relational frame skills training intervention to increase general intelligence and scholastic aptitude. *Learning and Individual Differences, 47*, 222–235.

De Houwer, J. (2014). A propositional model of implicit evaluation. *Social and Personality Psychology Compass, 8*(7), 342–353.

Dougher, M. J., Augustson, E., Markham, M. R., Greenway, D. E., & Wulfert, E. (1994). The transfer of respondent eliciting and extinction functions through stimulus equivalence classes. *Journal of Experimental Analysis of Behavior, 62*, 331–351.

Dougher, M. J., Hamilton, D. A., Fink, B. C., & Harrington, J. (2007). Transformation of the discriminative and eliciting functions of generalized relational stimuli. *Journal of Experimental Analysis of Behavior, 88*(2), 179–197. doi:10.1901/jeab.2007.45-05.

Dymond, S., & Barnes, D. (1995). A transformation of self-discrimination response functions in accordance with the arbitrarily applicable relations of sameness, more-than and less-than. *Journal of Experimental Analysis of Behavior, 84*, 163–184.

Dymond, S., & Roche, B. (2013). *Advances in relational frame theory: research and application*. CA: New Harbinger Publications, Inc..

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

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

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

Hayes, S. C., & Hayes, L. (1992). Verbal relations and the evolution of behavior analysis. *American Psychologist, 47*, 1383–1395.

Hayes, S. C., Kohlenberg, B. K., & Hayes, L. J. (1991). The transfer of specific and general consequential functions through simple and conditional equivalence classes. *Journal of Experimental Analysis of Behavior, 56*, 119–137.

Hayes, J., & Stewart, I. (2016). Comparing the effects of derived relational training and computer coding on intellectual potential in school-age children. *British Journal of Educational Psychology, 86*, 397–411. doi:10.1111/bjep.12114.

Healy, O., Barnes-Holmes, D., & Smeets, P. M. (2000). Derived relational responding as generalized operant behavior. *Journal of Experimental Analysis of Behavior, 74*(2), 207–227.

Hughes, S. & Barnes-Holmes, D. (2011). On the formation and persistence of implicit attitudes: New evidence from the implicit relational assessment procedure (IRAP). *The Psychological Record, 61*, 391–410.

Hughes, S., & Barnes-Holmes, D. (2013). Relational frame theory: the basic account. In S. Hayes, D. Barnes-Holmes, R. Zettle, & T. Biglan (Eds.), *The Wiley handbook of contextual behavioral science*. New York: Wiley-Blackwell.

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

Ju, W. C., & Hayes, S. C. (2008). Verbal establishing stimuli: testing the motivative effect of stimuli in a derived relation with consequences. *Psychological Record, 58*, 339–363.

Killeen, P. R., & Jacobs, K. W. (2017). Coal is not black, snow is not white, food is not a reinforcer. *The Behavior Analyst* . doi:10.1007/s40614-016-0080-7.

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

Pepper, J. W., & Herron, M. D. (2008). Does biology need an organism concept? *Biological Reviews, 83*(4), 621–627. doi:10.1111/j.1469-185X.2008.00057.

Plumb, J. C., Stewart, I., Dahl, J., & Lundgren, T. (2009). In search of meaning: values in modern clinical behavior analysis. *The Behavior Analyst, 32*, 85–103.

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

Nosek, B. A., Hawkins, C. B., & Frazier, R. S. (2011). Implicit social cognition: from measures to mechanisms. *Trends in Cognitive Sciences, 15*(4), 152–159.

O'Hora, D., Barnes-Holmes, D., Roche, B., & Smeets, P. (2004). Derived relational networks as novel instructions: a possible model of generative verbal control. *Psychological Record, 54*, 437–460.

O'Hora, D., Barnes-Holmes, D., & Stewart, I. (2014). A relational network model of derived instructional control: antecedent and consequential control of derived instruction following. *Journal of Experimental Analysis of Behavior, 102*(1), 66–85.

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

O'Hora, D., Peláez, M., Barnes-Holmes, D., Rae, G., Robinson, K., & Chaudhary, T. (2008). Temporal relations and intelligence: correlating relational performance with performance on the WAIS—III. *Psychological Record, 58*(4), 569–584.

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

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

Roche, B., & Barnes, D. (1997). The behavior of organisms? *Psychological Record, 47*, 597–618.

Sidman, M. (1971). Reading and auditory-visual equivalences. *Journal of Speech, Learning and Hearing Research, 14*, 5–13.

Sidman, M. & Tailby, W. (1982). Conditional discrimination vs. matching to sample: An expansion of the testing paradigm. *Journal of the Experimental Analysis of Behavior, 37*(1), 5–22.

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 Experimental Analysis of Behavior, 37*(1), 23–44. doi:10.1901/jeab.1982.37-23.

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

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

Stewart, I. (2016). The fruits of a functional approach for psychological science. *International Journal of Psychology, 51*(1), 15–27. doi:10.1002/ijop.12184.

Stewart, I., Barnes-Holmes, D., Roche, B., & Smeets, P. W. (2002). A functional-analytic model of analogy: a relational frame analysis. *Journal of Experimental Analysis of Behavior, 78*, 375–396. doi:10.1901/jeab.2002.78-375.

Stewart, I., Hooper, N., Walsh, P., O'Keefe, R., Joyce, R., & McHugh, L. (2015). Transformation of thought suppression functions via same and opposite relations. *Psychological Record, 65*(2), 375–399.

Stewart, I., & McElwee, J. (2009). Relational responding and conditional discrimination procedures: an apparent inconsistency and clarification. *The Behavior Analyst, 32*, 309–317.

Stewart, I., McElwee, J., & Ming, S. (2013). Language generativity, response generalization, and derived relational responding. *The Analysis of Verbal Behavior, 29*(1), 137–155.

Stewart, I., & Roche, B. (2013). Relational frame theory: an overview. In S. Dymond & B. Roche (Eds.), *Advances in relational frame theory: research and application*. Oakland, CA: New Harbinger Publications, Inc.

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.

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

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

Whelan, R., Barnes-Holmes, D., & Dymond, S. (2006). The transformation of consequential functions in accordance with the relational frames of more-than and less-than. *Journal of Experimental Analysis of Behavior, 86*(3), 317–335.

Zettle, R. D., Hayes, S. C., Barnes-Holmes, D., & Biglan, A. (2016). *The Wiley handbook of contextual behavioral science*. New York: Wiley-Blackwell.