Available online at www.sciencedirect.com

ScienceDirect

Current Opinion in Psychology

From Relational Frame Theory to implicit attitudes and back again: clarifying the link between RFT and IRAP research

Ian Hussey, Dermot Barnes-Holmes and Yvonne Barnes-Holmes

Relational Frame Theory (RFT) is a functional-analytic account of human language and cognition, including human psychopathology. The core premise of the theory is that language and cognition is composed of relational acts. Over the past 10 years, the theory has served to generate the development of a measure, known as the Implicit Relational Assessment Procedure, which was designed initially to provide a metric of the strength or persistence of relational responding. Although the IRAP provides a measure of implicit attitudes, we argue that it is time to refocus on its original purpose: measuring the strength of relational framing. This refocusing has already started to generate a new conceptual framework, which potentially will lead to improved functional specificity for behavior therapy.

Addresses
National University of Ireland Maynooth, Ireland

Corresponding authors: Hussey, Ian (ian.hussey@nuim.ie), Barnes-Holmes, Dermot (dermot.barnes-holmes@nuim.ie)

Current Opinion in Psychology 2015, **2**:11–15

This review comes from a themed issue on **Third wave behavioural therapies**

Edited by **Kevin Vowles**

http://dx.doi.org/10.1016/j.copsyc.2014.12.009

Introduction

Psychological science faces two enduring questions: *what to measure* and *how to measure it*. The current article will discuss the contributions of behavioral psychology to both of these questions over the last 15 years. Specifically, we will clarify the link between the ongoing development of a modern behavioral account of language and cognition, Relational Frame Theory (RFT) [1], and the use of a class of methodologies known as implicit measures [2], which have been used to study complex human behavior, including psychopathology (hereafter referred to as psychological suffering).

Relational Frame Theory

As discussed by Guinther and Dougher elsewhere in this special issue [3], behavioral psychology has generated a range of behavioral principles (e.g. reinforcement, punishment, extinction, stimulus generalization [4]) that have proven very useful in predicting-and-influencing the behavior of non-humans. However, it has been argued that these principles do not adequately account for the behavior of verbally able humans [1]. The development of RFT provided a conceptual account of the phenomena that define language and higher cognition in verbally able adult humans, including domains such as analogical reasoning, perspective taking, and psychological suffering [5].

At the time of writing the original RFT 'purple book', the theory was at a stage of development where it could provide a *conceptual account* of many clinically relevant phenomena that existing behavioral principles could not (e.g. developing a phobia of an animal that the individual had never directly encountered) [1]. Furthermore, it allowed researchers to *model* many of these behaviors within the laboratory (e.g. derived acquisition of spider fear via a respondent learning paradigm) [6]. However, our ability to *capture* verbal behavior (i.e. arbitrarily applicable relational responding) 'in-flight', as it is emitted by an individual, was relatively limited. Existing paradigms within behavioral psychology, such as matching-to-sample (MTS), provided a binary or dichotomous outcome. For example, an individual in a given study would either demonstrate the derived acquisition of fear of spiders (following training and testing with MTS) or he would not. This inevitably invited a binary or dichotomous way of thinking about relational framing itself. Rarely did we ask about the relative strength, probability, or persistence of relational responding.

In an effort to break out of dichotomous ways of thinking about relational framing, some RFT researchers began a concerted effort to develop a new paradigm that was capable of capturing such 'in-flight' verbal behaviors. This resulted in the development of the Implicit Relational Assessment Procedure, which has led to a behaviorally oriented research program on implicit cognition per se (IRAP) [7]. Nevertheless, it is important to emphasize that our interest in expanding the basic RFT account of language and cognition, above and beyond implicit cognition, never waned during this process. Hence, the title of this article employs the metaphor of a return journey back to RFT.

The development of implicit measures

Around the time the 'purple book' was being written, cognitive and social psychology developed and became

interested in a class of methodologies referred to as implicit measures. Instead of explicitly asking individuals using self-report measures what they think, feel, believe, or intend, these computer-based behavioral measures infer attitudes and response biases from reaction times or other behavioral metrics [8]. For example, the most well known implicit measure, the Implicit Association Test (IAT), requires individuals to pair images or words (e.g. white and black faces with positive and negative words) under time pressure and according to opposing rules across blocks of trials (e.g. block 1: black people — negative and white people — positive versus block 2: black people — positive and white people — negative) [9]. The average difference in milliseconds between the blocks represents a bias toward finding it 'easier' to respond on one block relative to the other, and is thought to reveal an 'implicit' attitude.

An explosion of research using implicit measures quickly found that they often produced unexpected results that diverged from those obtained via self-report. For example, individuals who expressed overtly egalitarian views were nonetheless found to show biases on the IAT related to race, religion, age, nationality, gender, and sexuality [9–13]. More importantly, implicit biases were shown to increase researchers' ability to predict future behavior above and beyond self-report measures. The basic argument is that implicit measures may be less sensitive to variables pertaining to social desirability and conscious awareness than standard self-reports, and thus account for increased variance. For example, the IAT has been shown to predict future voting patterns among individuals who explicitly report that they have not yet make up their minds [13], and also prediction of future self-harm and suicide attempts over and above existing self-report measures or personal or clinical estimations of risk [14,15].

The Implicit Relational Assessment Procedure

Certainly, measures such as the IAT provide impressive predictions of socially and clinically relevant behavior. Nonetheless, it is critical to note here that our interest in implicit cognition came out of our interest in understanding the dynamics of arbitrarily applicable relational responding *as it actually occurs*. To this end, and inspired in part by the IAT, the second author sought to develop a measure that was capable of producing data on the relative strength of individual relational responses. Nevertheless, the IRAP quickly emerged as a measure of implicit cognition, but one that differed from the IAT.

Put simply, the two primary differences between the IRAP and the IAT are as follows: the IAT produces a single metric of overall bias whereas the IRAP assesses individual relational responses, and the IRAP is unique in that it can assess the relative strength of semantically complex stimuli (i.e. relational networks over and above simple equivalence relations [16]). With regard to assessing individual relational responses, the IAT produces a single metric of overall bias, whereas the IRAP produces four metrics, one for each of the individual relational responses that are targeted by the measure (e.g. black people — positive, black people — negative, white people — positive, and white people — negative [17]). With regard to semantic complexity, the IAT is suited to assessing simple stimulus pairing, such as 'self' or 'others' with 'positive' or 'negative' [18], whereas the IRAP can be used to assess the relative strength of complex relations in, for example, 'actual' versus 'ideal' self-esteem (e.g. 'I am — valuable' versus 'I want to be — valuable' [19]).

Since the development of the IRAP, RFT researchers have pursued a course similar to that of researchers using other implicit measures, such as the IAT. The properties of the method have been explored and refined [20], the dissociation between results obtained using explicit (self-report) and implicit measures has been explored [21], and, most importantly in the context of the current special issue, an increasing number of studies have examined the IRAP's ability to shed light on a variety of clinically relevant phenomena [22]. For example, we have used the IRAP to detect attitudes related to the sexualization of children among sex offenders [23], predict of future cocaine relapses among individual in court-ordered rehabilitation programs [24•], provide the first empirical support of cognitive behavior therapy models of disgust within OCD [25], and distinguish between epileptic and psychogenic nonepileptic seizures [26].

Doubtless, such studies were pivotal in establishing the utility of the IRAP. Nonetheless, such research activities have, frankly, often served as somewhat of a distraction from the far more ambitious goal of advancing the RFT account of language, cognition, and psychological suffering itself. Recent conceptual developments that have emerged directly from the IRAP have, however, served to reorient us back to this broader scientific agenda.

Contributions from implicit measures to the ongoing development of RFT

The IRAP therefore provided a way to assess the relative strength of relational responding that was non-dichotomous. And, in developing and using the IRAP, so too did our thinking about relational responding become non-dichotomous. Whereas RFT had previously sought merely to demonstrate its basic units, such as relational frames, we had now reached a point where we were capable of asking questions about the dynamics and properties of that relational responding [27••]. To illustrate, rather than asking *whether* fear functions can be derived through a relational network, the research began to focus on questions concerning the extent to which fear responding *differs* across the first, the 100th or 1000th time it is derived. That is, we started to ask what determines whether a relational response is rapid or slow, or malleable

or rigid, thus connecting more directly with the rather vague concept of 'psychological flexibility' [28].

Several properties of relational responding have been elaborated, such as the extent to which a particular pattern of relational responding has occurred in the past ('derivation'), the various ways in which patterns of relational responding may differ in terms of properties such as number of stimuli, relations, and varieties of contextual control ('complexity'), and the extent to which a particular pattern of relational responding yields relatively consistent consequences ('coherence'). Certainly, these properties were first discussed in detail within the context of the IRAP and RFT's account of implicit attitudes, the Relational Elaboration and Coherence (REC) model [7,29]. However, our current and enduring focus goes far beyond a behavioral account of implicit cognition, to the construction of a framework for analyzing the operant units of arbitrarily applicable relational responding that characterize relatively simple to highly complex instances of human language and cognition, and thus psychological suffering. Specifically, recent conceptual advances have begun to consider how the three properties listed above, and others, such as fluency and levels of behavioral development, interact in a multi-dimensional multi-level manner [27••,29]. The following section will elaborate the importance of this scope, which may not be apparent at first glance.

Implications for behavior therapy

Existing research has already demonstrated that RFT allows us to conceptualize many forms of psychological suffering as a natural — if undesirable — product of our 'ability to language' (e.g. anxiety and phobia [6], derived transfer of mood functions [30]), rather than the necessary and/or sole product of a fundamental mental or biological pathology (although, it is important to note that this does not in any way exclude biological variables from our analysis [31,32]). Furthermore, RFT research has made many advances in modeling and unpacking the role of human language in psychological suffering within the laboratory (e.g. separating derived fear and avoidance [33••]; eliminating avoidance with therapeutic analogs [34••]; demonstrating derived thought suppression [35]; reactivity within depression [16]; disgust within OCD [25]; and ideal versus actual self-esteem within depression [19]).

We are hopeful that the ongoing development of RFT will serve to advance our ability to predict-and-influence behavior, particularly within the context of psychotherapeutic interventions. In our view, however, RFT currently specifies the core *behavioral process at work* (i.e. arbitrarily applicable relational responding [1]), but our understanding of the specific properties and dynamics of relational responding are not yet sufficiently advanced to the point where we can specify RFT-defined functional analytic therapeutic interventions (i.e. *operations that target specific properties of relational responding*). As an aside, we have recently argued that Acceptance and Commitment Therapy's 'middle-level terms' (e.g. defusion, self-as-context, etc. [36]) do not provide the level of functional specificity that we are seeking here [37•], as elaborated by McEnteggart and colleagues elsewhere in this special issue [38]. In fact, until this point is reached, we feel that there can be no legitimate claim to have created a so-called 'relational frame therapy'. In fact, the concept of 'relational frame therapy' is, to us, a misnomer, because behavioral principles may be applied equally in many contexts (e.g. education, psychotherapy, etc.) and within multiple approaches (e.g. CBT, ACT, DBT, etc.).

Conclusion

Some of the recent advances made within RFT research into psychological suffering have been a direct consequence of the development of an implicit measure and a conceptual analysis of that measure in terms of specific properties of relational responding (e.g. derivation, complexity, and coherence [27]). Nonetheless, our interest was never in implicit cognition per se, but instead we were seeking a deeper appreciation of the properties and dynamics of relational framing itself. We are hopeful that a richer understanding of relational framing, within a multi-dimensional multi-level framework [27••], will aid us in bringing functional precision and specificity to behavior therapy.

Conflict of interest

The authors report that they have no conflicts of interest.

References and recommended reading

Papers of particular interest, published within the period of review, have been highlighted as:

- • of special interest
- •• of outstanding interest

1. Hayes SC, Barnes-Holmes D, Roche B: *Relational Frame Theory: A Post-Skinnerian Account of Human Language and Cognition*. Plenum Press; 2001.
2. De Houwer J, Teige-Mocigemba S, Spruyt A, Moors A: **Implicit measures: a normative analysis and review**. *Psychol Bull* 2009, **135**:347-368.
3. Guinther PM, Dougher MJ: **The clinical relevance of stimulus equivalence and Relational Frame Theory in influencing the behavior of verbally competent adults**. *Curr Opin Psychol* 2015 http://dx.doi.org/10.1016/j.copsyc.2015.01.015. in this issue.
4. Catania AC: *Learning*. Prentice-Hall; 1998.
5. Hughes S, Barnes-Holmes D: **Relational Frame Theory: the basic account**. In *Handbook of Contextual Behavioral Science*. Edited by Hayes SC, Barnes-Holmes D, Zettle RD, Biglan A. Wiley-Blackwell; in press.
6. Smyth S, Barnes-Holmes D, Forsyth JP: **A derived transfer of simple discrimination and self-reported arousal functions in spider fearful and non-spider-fearful participants**. *J Exp Anal Behav* 2006, **85**:223-246.
7. Barnes-Holmes D, Barnes-Holmes Y, Stewart I, Boles S: **A sketch of the Implicit Relational Assessment Procedure IRAP and the**

Relational Elaboration and Coherence REC model. *Psychol Rec* 2010, **60**:527-542.

8. De Houwer J, Moors A: **How to define and examine the implicitness of implicit measures**. In *Implicit Measures of Attitudes: Procedures and Controversies.* Edited by Schwarz N, Wittenbrink B. Guilford Press; 2007:179-194.

9. Greenwald AG, McGhee DE, Schwartz JL: **Measuring individual differences in implicit cognition: the implicit association test**. *J Pers Soc Psychol* 1998, **74**:1464-1480.

10. Rudman LA, Greenwald AG, Mellott DS, Schwartz JLK: **Measuring the automatic components of prejudice: flexibility and generality of the implicit association test**. *Soc Cogn* 1999, **17**:437-465.

11. Rudman LA, Glick P: **Prescriptive gender stereotypes and backlash toward agentic women**. *J Soc Issues* 2001, **57**:743-762.

12. Jellison WA, McConnell AR, Gabriel S: **Implicit and explicit measures of sexual orientation attitudes: in group preferences and related behaviors and beliefs among gay and straight men**. *Pers Soc Psychol Bull* 2004, **30**:629-642.

13. Greenwald AG, Smith CT, Sriram N, Bar-Anan Y, Nosek BA: **Implicit race attitudes predicted vote in the 2008 US presidential election**. *Anal Soc Issues Public Policy* 2009, **9**:241-253.

14. Nock MK, Park JM, Finn CT, Deliberto TL, Dour HJ, Banaji MR: **Measuring the suicidal mind: implicit cognition predicts suicidal behavior**. *Psychol Sci* 2010, **21**:511-517.

15. Randall JR, Rowe BH, Dong KA, Nock MK, Colman I: **Assessment of self-harm risk using implicit thoughts**. *Psychol Assess* 2013, **25**:714-721.

16. Hussey I, Barnes-Holmes D: **The implicit relational assessment procedure as a measure of implicit depression and the role of psychological flexibility**. *Cogn Behav Pract* 2012, **19**:573-582.

17. Barnes-Holmes D, Murphy A, Barnes-Holmes Y, Stewart I: **The Implicit Relational Assessment Procedure: exploring the impact of private versus public contexts and the response latency criterion on pro-white and anti-black stereotyping among white Irish individuals**. *Psychol Rec* 2010, **60**:57-66.

18. Gemar MC, Segal ZV, Sagrati S, Kennedy SJ: **Mood-induced changes on the Implicit Association Test in recovered depressed patients**. *J Abnorm Psychol* 2001, **110**:282-289.

19. Remue J, De Houwer J, Barnes-Holmes D, Vanderhasselt MA, De Raedt R: **Self-esteem revisited: performance on the implicit relational assessment procedure as a measure of self- versus ideal self-related cognitions in dysphoria**. *Cogn Emot* 2013, **27**:1441-1449.

20. McKenna IM, Barnes-Holmes D, Barnes-Holmes Y, Stewart I: **Testing the fake-ability of the Implicit Relational Assessment Procedure IRAP: the first study**. *Int J Psychol Psychol Ther* 2007, **7**:253-268.

21. Power P, Barnes-Holmes D, Barnes-Holmes Y, Stewart I: **The Implicit Relational Assessment Procedure IRAP as a measure of implicit relative preferences: a first study**. *Psychol Rec* 2009, **59**:621-640.

22. Vahey N, Nicholson E, Barnes-Holmes D: **A meta-analysis of criterion effects for the Implicit Relational Assessment Procedure (IRAP) in the clinical domain**. *J Behav Ther Exp Psy*; in press.

23. Dawson DL, Barnes-Holmes D, Gresswell DM, Hart AJ, Gore NJ: **Assessing the implicit beliefs of sexual offenders using the implicit relational assessment procedure: a first study sex abuse**. *J Res Treat* 2009, **21**:57-75.

24. • Carpenter KM, Martinez D, Vadhan NP, Barnes-Holmes D, Nunes EV: **Measures of attentional bias and relational responding are associated with behavioral treatment outcome for cocaine dependence**. *Am J Drug Alcohol Abuse* 2012, **38**:146-154.

This study used the IRAP to predict cocaine abstinence within a court-ordered treatment program. The IRAP was shown to be a better predictor of abstinence than self-report measures, and initially more predictive than an attentional bias task.

25. Nicholson E, Barnes-Holmes D: **Developing an implicit measure of disgust propensity and disgust sensitivity: examining the role of implicit disgust propensity and sensitivity in obsessive–compulsive tendencies**. *J Behav Ther Exp Psychiatry* 2012, **43**:922-930.

26. Dimaro L, Dawson DL, Roberts NA, Brown I, Moghaddam NG, Reuber M: **Anxiety and avoidance in psychogenic nonepileptic seizures: the role of implicit and explicit anxiety**. *Epilepsy Behav* 2014, **33**:77-86.

27. •• Barnes-Holmes D, Barnes-Holmes Y, Hussey I: **Relational Frame Theory: finding its historical and intellectual roots and reflecting upon its future development**. In *Handbook of Contextual Behavioral Science*. Edited by Hayes SC, Barnes-Holmes D, Zettle RD, Biglan A. Wiley-Blackwell; in press.

This chapter contains a historical commentary of the RFT research agenda, and outlines the proposal that future work should focus on the interactions among various properties of relational responding within a multi-dimensional, multi-level framework.

28. Bond FW, Hayes SC, Baer RA, Carpenter KM, Guenole N, Orcutt HK, Waltz T, Zettle RD: **Preliminary psychometric properties of the Acceptance and Action Questionnaire-II: a revised measure of psychological inflexibility and experiential avoidance**. *Behav Ther* 2011, **42**:676-688.

29. Hughes S, Barnes-Holmes D, Vahey N: **Holding on to our functional roots when exploring new intellectual islands: a voyage through implicit cognition research**. *J Context Behav Sci* 2012, **1**:17-38.

30. Cahill J, Barnes-Holmes Y, Barnes-Holmes D, Rodriguez-Valverde M, Luciano C, Smeets PM: **The derived transfer and reversal of mood functions through equivalence relations: II**. *Psychol Rec* 2007, **57**:373-389.

31. Barnes-Holmes D, Regan D, Barnes-Holmes Y, Commins S, Walsh D, Stewart I, Smeets PM, Whelan R, Dymond S: **Relating derived relations as a model of analogical reasoning: reaction times and event-related potentials**. *J Exp Anal Behav* 2005, **84**:435-451.

32. Cochrane A, Barnes-Holmes D, Barnes-Holmes Y, Stewart I, Luciano C: **Experiential avoidance and aversive visual images: response delays and event-related potentials on a simple matching task**. *Behav Res Ther* 2007, **45**:1379-1388.

33. •• Luciano C, Valdivia-Salas S, Ruiz FJ, Rodríguez-Valverde M, Barnes-Holmes D, Dougher MJ, Cabello F, Sánchez V, Barnes-Holmes Y, Gutierrez-Martínez O: **Extinction of aversive eliciting functions as an analog of exposure to conditioned fear: does it alter avoidance responding?** *J Context Behav Sci* 2013, **2**:120-134.

This study substantiates, for the first time, two of the core premises of the Acceptance and Commitment Therapy model. Specifically, that emotions (i.e. fear) and avoidance are two functionally distinct behavioral repertoires, and that pathological reactions to stimuli are defined by persistent avoidance and not persistent emotions.

34. •• Luciano C, Valdivia-Salas S, Ruiz FJ, Rodríguez-Valverde M, Barnes-Holmes D, Dougher MJ, Cabello F, Sánchez V, Barnes-Holmes Y, Gutierrez-Martínez O: **Effects of an acceptance/defusion intervention on experimentally induced generalized avoidance: a laboratory demonstration**. *J Exp Anal Behav* 2014, **101**:94-111.

This analog study demonstrates, for the first time, the independent manipulation of avoidance functions and emotional (i.e. fear) functions. ACT-informed interventions were shown to effectively undermine (pathological) avoidance without altering (non-pathological) autonomic fear responding. It therefore provides evidence for Acceptance and Commitment Therapy's proposed method of action specifically, rather than merely its efficacy generally.

35. Hooper N, Saunders J, McHugh L: **The derived generalization of thought suppression**. *Learn Behav* 2010, **38**:160-168.

36. Hayes SC, Strosahl K, Wilson KG: *Acceptance and Commitment Therapy: An Experiential Approach to Behavior Change*. Guilford Press; 1999.

37. • Barnes-Holmes Y, Hussey I, McEnteggart C, Barnes-Holmes D, Foody M: **Scientific ambition: the relationship between Relational Frame Theory and middle-level terms in Acceptance and Commitment Therapy**. In *Handbook of Contextual Behavioral Science*. Edited by Hayes SC, Barnes-Holmes D, Zettle RD, Biglan A. Blackwell-Wiley; in press.
This chapter critically assesses the functionality and scientific utility of Acceptance and Commitment Therapy's 'middle level terms' (e.g. defusion), discusses the relationship between basic science and applied work, and argues for a refocusing on basic research

38. McEnteggart C, Barnes-Holmes Y, Hussey I, Barnes-Holmes D: **Ties between basic science of language and cognition to clinical applications**. *Curr Opin Psychol* 2015 http://dx.doi.org/10.1016/j.copsyc.2014.11.017. in this issue.