Relational Frame Theory:
Some Implications for Understanding and Treating Human Psychopathology

Yvonne Barnes-Holmes[*1], Dermot Barnes-Holmes[1], Louise McHugh[1],

and Steven C. Hayes[2]

[1]*National University of Ireland, Maynooth, Ireland,* [2]*University of Nevada, Reno, USA*

ABSTRACT

In the current paper, we attempt to show how both the basic and applied sciences of behavior analysis have been transformed by the modern research agenda in human language and cognition, known as Relational Frame Theory (RFT). At the level of basic process, the paper argues that the burgeoning literature on derived stimulus relations calls for a reinterpretation of complex human behavior that extends beyond a purely contingency-based analysis. Specifically, the paper aims to show how a more complete account of complex human behavior includes an analysis of relational frames, relational networks, relating relations, rules, perspective-taking, and the concept of self. According to the theory, this analysis gives rise to a new interpretation of human psychopathology that necessarily transforms the applied science of behavior therapy. The current paper is divided into three parts. In Part 1, we provide a brief summary of the integrated history of behavioral psychology and behavior therapy, including their emphases on the principles of classical and operant conditioning as the basis for an account of human psychopathology. In Part 2, the core features of RFT are presented, including the three concepts of bidirectional stimulus relations, relating relations, and rule-governance that constitute critical components of the RFT approach to human psychopathology. The paper therein attempts to illustrate, with the use of clinically relevant examples, the ways in which these concepts can be used to understand psychopathology and psychotherapy. In Part 3, RFT interpretations of three central features of Acceptance and Commitment Therapy (ACT), namely acceptance, defusion, and values are provided with a view to demonstrating the utility of basic RFT concepts in the treatment of human suffering.
Key Words: Experimental and applied behavior analysis, RFT, language and cognition, derived stimulus relations, complex human behavior.

RESUMEN

El presente trabajo pretende mostrar cómo las ciencias básica y aplicada del análisis de conducta han sido transformadas por el moderno programa de investigación acerca del lenguaje y la cognición denominado Teoría del Marco Relacional (RFT). A nivel básico, el artículo argumenta que la floreciente literatura sobre relaciones derivadas entre estímu-

* Correspondence should be directed to: Yvonne Barnes-Holmes at the Department of Psychology, National University of Ireland, Maynooth, Maynooth, Co. Kildare, Ireland. (Email: yvonne.barnes-holmes@may.ie).

los clama por una reinterpretación del comportamiento humano complejo más allá del mero análisis basado en contingencias. Concretamente, este artículo se propone mostrar cómo una explicación más completa del comportamiento humano complejo incluye un análisis de los marcos relacionales, las redes de relaciones, las relaciones entre relaciones, las reglas, la toma de perspectiva y el concepto del yo. De acuerdo con esta teoría, este análisis da lugar a una nueva interpretación de la psicopatología que necesariamente supone una transformación de de la ciencia aplicada de la terapia de conducta. El presente artículo se divide en tres partes. En la primera se presenta un breve resumen de la historia integrada de la psicología conductual y la terapia de conducta, incluyendo el énfasis de éstas en los principios del condicionamiento, clásico y operante, como base explicativa de la psicopatología humana. En la segunda parte se presentan las características funda-mentales de la RFT, incluyendo los tres conceptos de relaciones bidireccionales entre estímulos, relaciones entre relaciones y seguimiento de regla, que constituyen los com-ponentes críticos de la aproximación de la RFT a la psicopatología humana. A partir de ahí, el artículo intenta ilustrar, con el uso de ejemplos clínicamente relevantes, de qué modo estos conceptos pueden ser utilizados para entender la psicopatología y la psicote-rapia. En la tercera y última parte se dan interpretaciones basadas en la RFT de tres características centrales de la Terapia de Aceptación y Compromiso (ACT) como son la aceptación, la defusión cognitiva y los valores, con objeto de demostrar la utilidad de los conceptos básicos de la RFT para el tratamiento del sufrimiento humano.

Palabras clave: Análisis experimental y aplicado del comportamiento, RFT, lenguaje y cognición, relaciones derivadas entre estímulos, comportamiento humano complejo,

Researchers and clinicians, from a variety of psychological perspectives, assume that incidences or clusters of psychopathology involve abnormal behavior (Davison, Neale, & Krung, 2004). A number of key classification systems have been traditionally employed as a means of organizing these behaviors -and the thoughts, feelings, and other internal states to which they are related- into syndromes or functional classes that can readily be discussed in the scientific and lay communities. The most widely used categorical model of classification is DSM (the Diagnostic and Statistical Manual) in which clusters of abnormal behaviors, thoughts, and feelings are grouped as syndromes, that in turn are organized across five interrelated classification axes (American Psychiatric Association, 1994). From a strictly behavioral perspective, however, syndromal classification appears to be of limited utility in identifying common *functional* dimensions of abnormal behavior (Wilson, Hayes, Gregg, & Zettle, 2001). It is perhaps for this reason that syndromal classification has not been adopted as a primary diagnostic tool by researchers and practitioners who advocate a functional, and typically behavior-analytic, approach to psychological problems (Hayes, Wilson, Gifford, Follette, & Strosahl, 1996). In contrast, these professionals demonstrate a preference for a dimensional model of classification that emphasizes functional behavioral overlaps or commonalities that directly guide assessment and treatment (e.g., the presence of high levels of emotional avoidance, see Hayes, Nelson, & Jarrett, 1987). Identification of these commonalities is driven almost entirely by a core set of basic behavioral principles, thereby reflecting a more integrated scientist-practitioner approach to psychopathology than that offered by syndromal classifications (Barlow, Hayes, & Nelson, 1985).

One important corollary of the functional or behavior-analytic approach to abnormal behavior, that arises from its reliance upon a core set of basic behavioral principles, is the need to effect appropriate changes in the applied science when the basic scientific principles are modified or extended. The current paper argues that such a change is presently underway in the science of behavior analysis under the rubric of Relational Frame Theory (RFT, Hayes, Barnes-Holmes, & Roche, 2001). Specifically, these researchers have argued that the burgeoning literature on derived stimulus relations calls for a reinterpretation of complex human behavior, including most types of abnormal behavior, that extends beyond a purely contingency-based analysis. The primary aim of the current paper is to describe the behavioral account of human language and cognition, known as RFT, and its implications for a new functional interpretation of human psychopathology as complex human verbal behavior.

The current article is divided into three parts. In Part 1, we provide a brief summary of the integrated history of behavioral psychology and behavior therapy, including their emphases on the principles of classical and operant conditioning as the basis for an account of human psychopathology. In Part 2, the core features of RFT are presented, including the three concepts of bidirectional stimulus relations, relating relations, and rule-governance that constitute critical components of the RFT approach to human suffering. The paper therein attempts to illustrate, with the use of clinically relevant examples, the ways in which these concepts can be used to understand psychopathology and psychotherapy. In Part 3, RFT interpretations of three central features of Acceptance and Commitment Therapy (ACT), namely acceptance, defusion, and values are provided with a view to demonstrating the utility of basic RFT concepts in the treatment of human suffering.

PART 1
THE RELATIONSHIP BETWEEN BASIC PRINCIPLES AND TREATMENT

Early behavior therapists hailed 'Learning Theory' as the scientific basis from which to understand human psychopathology, and substantive empirical evidence of basic behavioral principles did much to stimulate the development of new intervention procedures and to generate clinical interest in the basic science (Mahoney, 1994). The focus on data-driven practice distinguished behavior therapy from the more traditional types of psychotherapy that were directed almost entirely by theoretical models (e.g., the psychodynamic interpretation of developmental conflicts as a source of psychological distress in later life –Malan, 1995). This integration between basic behavioral principles and therapy rested upon two key theoretical assumptions. First, the same principles observed with nonhuman populations operated in the context of human behavior (i.e., the continuity assumption –Hayes, Fox, Gifford, Wilson, Barnes-Holmes, & Healy, 2001; Skinner, 1938). Second, normal and abnormal human behavior involved the same basic behavioral processes (Wilson, et al., 2001).

The key processes relied upon most heavily in the behavior therapy movement were those identified as classical and operant conditioning. For example, in the now famous Little Albert experiment, Watson and Rayner (1920) reported evidence of

empirically-induced conditioned fear responses to neutral stimuli, and suggested this type of classical conditioning as a possible analog of the development of fears or phobias. Subsequently, Skinner's work on principles of reinforcement provided experimental analogs of the gradual shaping of even extreme forms of behavior, such as psychotic episodes (i.e., what he referred to as "negative utility") and its maintenance by operant contingencies (Wilson & Davidson, 1971). In broad terms, the growing body of empirical evidence for classical and operant conditioning processes seemed to allow the behavior therapy movement to circumvent the need to postulate abnormal internal events or indeterminable historical sources of causation.

With these theoretical advances, came more sophisticated and successful treatment regimes and concomitant increases in clinical popularity for behavior therapy in the 50's and 60's. Wolpe, for example, employed systematic desensitization and counter-conditioning in the treatment of anxiety and phobias and reported unprecedented therapeutic outcomes using behavioral methodologies (e.g., 90% cure or substantive improvement compared to the 66% recovery rate suggested by Eysenck; see Wolpe, 1952a, 1954, 1958). In addition to Wolpe's use of established behavioral principles, such as conditioning, several other researchers postulated new concepts or principles that had a significant impact on the behavior therapy movement. For example, Seligman (1974) provided a "learned helplessness" account of depression, and Bandura and Menlove's (1968) concept of modeling provided a possible explanation for the development of substance abuse or phobia.

Although clinicians were developing successful behavioral treatment regimes, the very foundation of the science itself, vis-à-vis human psychology, was being questioned, most notably by Chomsky (1959). Chomsky's vehement criticisms of Skinner's *Verbal Behavior* (1957) began to cast serious doubt on a behavioral account of complex psychological phenomena such as human language, and of course, of psychopathology. Coupled with increasing disenchantment with the behavioral approach in general, the stage was set for the so-called cognitive revolution of the 60's and 70's, and the subsequent emergence of cognitive therapy (see Moorey, 2003).

Some researchers, such as Bandura (1977), and Mahoney and Arnkoff (1978) attempted to offer integrative models that combined both cognition and behavior. Mahoney, for example, emphasized the role of cognitive processes such as expectation and attribution in conditioning. Similarly, Ellis (1958) highlighted the importance of irrational beliefs in behavioral problems, and harnessed this view in the development of Rational Emotive Therapy (RET). Although the work of practitioners such as Ellis did much to sustain the credibility and maintenance of the behavior therapy movement, there continued to be little or no change in the basic behavioral science in a manner that effectively addressed Chomsky's criticisms (Barnes-Holmes, Barnes-Holmes, Roche, Healy, Lyddy, Cullinan, & Hayes, 2001). However, the study of rule-governed behavior (see Hayes, 1989 for a review) and equivalence relations (see Sidman, 1994 for a review), and more recently the attempt to account for these two phenomena, and human language and cognition more generally, in terms of RFT (see Hayes, et al., 2001 for a review), provide the beginnings of an essential change in the basic science of behavior analysis.

The most common criticism of behavior analysis is that the basic principles, identified primarily with nonhuman organisms, cannot account for the *generativity* or *complexity* of human language and cognition (Barnes-Holmes, et al., 2001. Relational frame theorists appear to be in agreement with this view (Barnes-Holmes, Dymond, Grey, & Roche, 1999). The traditional behavioral focus on rats and pigeons was based on the continuity assumption that the basic scientific principles identified with these organisms were generally applicable to humans. This continuity assumption served its purpose well in the early days of behavioral psychology, and proved particularly useful in the development of the treatment regimes collectively described as Applied Behavior Analysis (e.g., Lovaas, 1981). According to RFT, however, the continuity assumption has served its time, and a new research agenda focused on human behavior per se is necessary (Barnes-Holmes, et al., 1999). In the following sections, we outline some of the core behavioral concepts that have driven this agenda, and thereafter consider their applicability to a behavioral interpretation of human psychopathology.

Bidirectional Stimulus Relations

From the perspective of RFT, the most basic difference between nonverbal and verbal behavior and between nonhumans and humans can be demonstrated by the following example of classical conditioning. Consider a pet dog who becomes excited at the sound of the word "walkies" because in the past this word predicts exercise. With this type of training history, the important psychological functions of exercise have become attached to a previously neutral event (e.g., the word "walkies"). Now consider what would happen if the order of events had been reversed by saying "walkies" only *after* the dog had been taken out. Evidence suggests that the dog will show little or no excitement when it hears the word "walkies" because this would require what has been called backward associative conditioning (Catania, 1998). Although animals can readily learn about neutral events (e.g., words) that *predict* the onset of more important psychological events (e.g., exercise), they do not readily learn about neutral events that *follow* important events (see Hall, 1984).

By contrast, a large body of empirical evidence paints a very different picture for humans. Imagine that a young girl who enjoys playing with a new toy is told *afterwards* that the toy was called "Nimbo". If, on a subsequent occasion, she is asked if she wants to play with Nimbo, the child may well show signs of excitement. In other words, the sound of the word "Nimbo" appears to make the child think of the new toy, even though the specific word has never *predicted* access to the toy. Unlike animals, therefore, humans can readily relate neutral events to more important psychological events, even when the former consistently follow the latter (see Leader, Barnes, & Smeets, 1996).

According to RFT, the bidirectionality of stimulus relations is a defining feature of human language and cognition, and makes a significant contribution to a behavior-

analytic account of private events (Hayes & Toarmino, 1999). For example, if human language was an entirely unidirectional (rather than bidirectional) process, it would be difficult to explain how individuals with psychological problems often avoid talking about psychologically painful events (because the talk would not predict the occurrence of the event in the future). It is not difficult to train a pigeon to produce a "self-report" on whether or not it has just been exposed to shock (i.e., simply reinforce responses on one key following shock, and on another following no shock). However, because the "self-report" in this case follows the shock it does not have the aversive functions of that event. In other words, the shock is aversive, but the report of it is not.

If the derived relations among events and the words that are used to describe them are bidirectional, as is the case with humans, the situation is very different. Reporting traumatic events, for instance, is difficult because the bidirectional relations between the words and the events allow the report to acquire many of the aversive and painful functions of the trauma itself (Hayes & Gifford, 1997).

Relational Networks

Another important feature of RFT's account of language and cognition describes the organization of stimulus relations into *relational networks*. Consider the young girl from the previous example who is told that her toy Nimbo has a secret name, "Daxy". Subsequently, whenever she hears the word "Daxy" she may think of the word "Nimbo" as well as the actual toy, even though the word "Daxy" has not yet been directly associated with the toy. This effect has been described as the emergence of *equivalence relations* among the actual toy and the words "Nimbo" and "Daxy", and numerous studies have demonstrated the ability of young children to learn large and complex networks of equivalence relations (e.g., Smeets, Barnes, & Roche, 1997; see also Luciano, Barnes-Holmes, & Barnes-Holmes, 2001).

According to RFT, relational networks containing stimulus relations, such as those of equivalence, provide the basis for a functional analysis of many aspects of human language and cognition. For example, if you are told that an arbitrary stimulus referred to as A is equal to another arbitrary stimulus B, and that B is equal to a third arbitrary stimulus C, then you will be able to determine the transitive relation of equivalence between A and C (i.e., A and C are equal). In this example, a very simple relational network containing equivalence relations among A, B, and C is constructed. Furthermore, if you were then told that A was tall, you could determine, without additional information, that B and C were also tall, by virtue of the fact that they are equal. These facts could be derived without ever having direct contact with the actual B or C stimuli.

Relational networks appear to overlap considerably with many natural language phenomena, such as naming and reading (e.g., deRose, deSouza, Rossito, & deRose, 1992; Sidman, 1971). For example, given training in the spoken word "cookie" and an actual cookie, and between the written word COOKIE and the spoken word "cookie," a child will identify the written word COOKIE as in an equivalence relation with "cookie" even though this performance has never actually been trained. Thus, symmetry

and transitivity between written words, spoken words, and objects is commonplace in naming and reading behaviors (Hayes, Gifford, & Ruckstuhl, 1996).

Although the concept of the relational network appears to be useful in *describing* at least some of the generativity of human language, it does not constitute a behavior-analytic account of these phenomena. In order to *explain* language and cognition (e.g., the emergence of derived relations between written and spoken words), RFT appeals to already-established behavioral principles (Hayes & Hayes, 1989).

Explaining Derived Relational Responding

It is a well-established fact that organisms respond readily to the formal relations among stimuli. For example, even insects have demonstrated the discrimination of the "dimmest" of an array of illuminated stimuli (Reese, 1968). This type of relational responding appears to be controlled primarily by formal or *nonarbitrary* stimulus relations (i.e., one of the stimuli is actually the dimmest). In addition, however, humans readily demonstrate patterns of relational responding that are not controlled only by the formal properties of the related events, but by specific contextual cues (García & Benjumea, 2001).

Consider the following example of contextual control by the simple word "is". During early natural language interactions, children are often presented with objects (e.g., a teddy) and are asked to repeat the object's name ("teddy"). This interaction may be described as 'see object X→ hear name Y→ repeat name Y'. At the same time, children are also taught the reverse sequence of events in which they are asked to identify objects upon hearing the name. This interaction may be described as 'hear name Y→ pick object X'. During early naming interactions, many specific examples of object-word and word-object relations are explicitly trained. According to RFT, when a child has been exposed to sufficient exemplars of responding to both types of relations, a repertoire of derived object-name and name-object relational responding is established (Barnes-Holmes, Barnes-Holmes, Roche, & Smeets, 2001a; Barnes-Holmes, Barnes-Holmes, Roche & Smeets, 2001b).

If a child with this type of naming history is instructed: "This is grandma," contextual cues (such as the word "is", and the general social context) predict that if this person is "grandma" (object X-name Y), then "grandma" is this person (name Y-object X). Consequently, the child may now identify the appropriate person when asked "Where is grandma?" in the absence of explicit reinforcement. According to RFT, although this type of relational response is derived (because there is no history of explicit reinforcement for pointing to grandma), it is not genuinely novel. This is described as a type of *generalized operant behavior* that has been brought under the control of contextual cues (e.g., the word "is") through a process of differential reinforcement.

In more abstract terms, imagine that exemplar training of a particular stimulus relation is provided; for example, A is B and B is A are both reinforced; C is D and D is C are both reinforced; E is F and F is E are both reinforced, and so on across additional exemplars. Following this training, according to RFT, a derived relational

response may now occur for a novel exemplar (e.g., if "X is Y" is reinforced, then "Y is X" is derived in the absence of differential reinforcement). In effect, operant contingencies select a particular pattern of relational responding in the presence of a specific contextual cue, and these contingencies are applied across numerous exemplars. As a result, the relational responding may generalize to other novel exemplars in the presence of the appropriate contextual cue, and thus, according to RFT, this performance constitutes an example of a generalized operant class (Barnes-Holmes & Barnes-Holmes, 2000).

According to RFT, other types of stimulus relations, including more-than and less-than; different from; and opposite to, that are prevalent in human language, also may be explained as generalized operant behavior. Consider a young child who is taught to select the larger of two cups of juice in response to the question: "Which cup has more?" and the smaller of two boxes in response to "Which has less?" In this case, the appropriate response is determined in part by the nonarbitrary relationship of physical size between the related objects. However, with appropriate exemplar training this relational response may be brought under the control of contextual cues that are purely conventional rather than formal or nonarbitrary. If a child is taught, for example, that "if A is more than B, then B is less than A" and "if C is more than D, then D is less than C" and so on across other exemplars, simply telling the child that "X is more than Y" may generate the derived relational response "Y is less than X". In this case, the relational response comes under the control of the words "more" or "less", rather than a formal stimulus dimension (Barnes & Roche, 1996; Barnes-Holmes & Barnes-Holmes, 2000). When this occurs, the relational response can now be *arbitrarily applied* to a range of other stimuli, even when their nonarbitrary properties (e.g., actual size) do not occasion the relational response. In Europe, for instance, a one-euro coin is described as worth *more than* a fifty-cent coin, even though the former is physically smaller than the latter. For RFT, deriving the appropriate relations in this case is an example of *arbitrarily applicable relational responding*, because the more-than and less-than relations have been arbitrarily applied to the coins by the verbal community (and run counter to the formal dimensions of the actual stimuli). This illustration provides yet another example of the way in which RFT explains some of the complex phenomena of language and cognition (e.g., understanding monetary value) in terms of a history of differential reinforcement that is then *generalized* to arbitrary, novel stimuli and events.

Relational Frames

Relational Frame Theory attempts to explain the richness of human language and cognition, as well as their generativity, and employs the concept of the *relational frame* in this regard. Put simply, relational frames describe the various patterns of derived relational responding that emerge, and include frames of coordination (including equivalence); comparison (including more-than and less-than); distinction; opposition; and hierarchy (Hayes, et al., 2001). For example, if you are told that 'A is the opposite of B and B is the opposite of C', then you will readily derive relations of opposition between B and A and between C and B, but a relation of sameness between A and C

and between C and A. Although relational frames are distinguished from one another, they share three defining properties: mutual entailment, combinatorial entailment, and the transformation of stimulus function.

Mutual entailment describes the relations between two stimuli or events. For example, if you are told that A is the same as B, then you will readily derive the mutually entailed relation of B as the same as A. *Combinatorial entailment* describes the relations among three or more stimuli. For example, if you are told that A is more than B and B is more than C, then you will readily derive that A is more than C and that C is less than A. Combinatorially entailed relations differ from mutually entailed relations not only in terms of the number of related events, but also in terms of specificity. For example, if A is more than B and A is more than C, then the entailed relations between B and C are unspecified (i.e., they cannot be determined –B and C may be the same, or one may be worth more/less than the other).

The *transformation of function* is the third defining feature of a relational frame and, according to RFT, this concept provides the psychological content for derived relations. To understand transformation of function at its simplest, imagine if you are told that 'A is more than B' and a reinforcing function is then attached to B (e.g., by pairing it with access to reinforcers). It is likely thereafter that A will acquire an even greater reinforcing function, in the absence of explicit training, because of its participation in a more-than relation with B (Dymond & Barnes, 1995; Roche & Barnes, 1997; Roche, Barnes-Holmes, Smeets, Barnes-Holmes, & McGeady, 2000; Visdómine & Luciano, 2002).

As is the case with mutually and combinatorially entailed relations, the transformations of function are also under contextual control (Barnes, Browne, Smeets, & Roche, 1995; Wulfert & Hayes, 1988). When asked to "Think of a glass of cold orange juice", you might begin to experience some of the perceptual functions of orange juice, such as cool and orange, even though no juice is actually present. In the language of RFT, the phrase "orange juice" participates in a frame of coordination with actual juice. The words "Think of a" provide a specific context in which visual functions of orange juice in particular are selected. For example, if you had otherwise been instructed to "Think of the smell of cold orange juice", this instruction may provide a context in which olfactory functions (e.g., sweetness) rather than visual functions would be elicited. Contextual cues, therefore, control the type of relational frame as well as the psychological functions that are transformed in accordance with it.

The concept of the transformation of function is important for understanding human psychopathology, particularly when combined with existing behavioral principles, such as classical conditioning. Imagine a young boy who is told that he is going to the "doctor", and is thereafter exposed to an uncomfortable medical examination (i.e., aversive functions of the word "doctor" are classically conditioned). Imagine now that the child is told that dentists are like doctors who look after your teeth (i.e., a relation of coordination or equivalence is established between doctor and dentist). When subsequently required to go to the dentist, the child will likely show signs of anxiety even though he has never actually had an aversive experience at the dentist. In this case, the aversive functions of doctor have transferred to dentist via the relations of

coordination. In effect, therefore, the child does not need to experience potentially aversive consequences of visiting the dentist in order to demonstrate feelings of anxiety.

Of course, functions may transform in accordance with derived relations other than equivalence and such effects may also help to explain certain aspects of psychopathology. Imagine a man who experiences an extreme panic attack in a small local store, and subsequently becomes agorophobic. During treatment, the man is asked to rank potential anxiety-provoking locations, and paradoxically he ranks the local store in which he experienced his first panic attack as *less* anxiety-provoking than Macy's department store in New York, which in fact he has never visited. According to a traditional classical conditioning perspective, the stimulus that preceded the panic attack –in this case the local store– should predict greater arousal than a previously unexperienced stimulus, which in principle differs along many stimulus dimensions from the conditional stimulus. According to RFT, however, Macy's likely participates in a number of comparative frames with the local store, such as: much larger than; more busy than; more dangerous than; less friendly than, and so on. Thus, the functions of Macy's are transformed in accordance with these multiple stimulus relations based on the directly established aversive psychological functions for the local store. In more colloquial terms, the client might say: "If I panicked in a small local store, imagine what would happen to me in a really large and busy department store in New York City."

Other researchers have similarly drawn on the concept of derived relations, and RFT more generally, in attempting to explain various aspects of human psychopathology, including anxiety (Friman, Hayes, & Wilson, 1998), depression (Hayes & Wilson, 1993), self esteem (Barnes, Lalor, Smeets, & Roche, 1996; Dymond & Barnes, 1995), and sexual dysfunction (Barnes & Roche, 1997b; Roche & Barnes, 1997, 1998). Our more immediate concern here, however, is to show how the basic concept of the derived relation may be extended to accommodate even more complex examples of human language and cognition than discussed thus far, and their implications for human suffering.

Relating Relations

According to RFT, relational networks can be related to other relational networks, and thus relations among stimuli that participate in and between relational networks can be extremely complex (Stewart, Barnes-Holmes, Hayes, & Lipkens, 2001). In fact, RFT employs the concept of relating relations in its analysis of metaphor, which appears to have some bearing on the understanding and treatment of human psychopathology (Wilson, Hayes, Gregg, & Zettle, 2001). For illustrative purposes, consider the example of the clinical metaphor: "Struggling with anxiety is like struggling in quicksand". Of course, the metaphor is employed to demonstrate to clients that struggling with anxiety makes anxiety worse in the same way as struggling in quicksand makes you sink more quickly. In the language of RFT, the relational network created by this metaphor may be conceptualized as follows: anxiety (A) is to psychological struggle (B) as quicksand (C) is to physical struggle (D). In this case, deriving the relation (e.g., of coordination) between anxiety and quicksand may permit the client to see that struggling with either leads to a formally similar outcome. The functions of anxiety, therefore, are transformed

for the client, such that s/he may now derive new cause and effect relations, such as "struggling with anxiety will only make it worse".

In the same way that relating relations or metaphor may be used in a therapeutic context as a means of changing important psychological functions, it may also be used to understand how human language and cognition feed into psychopathology in rather unusual ways. For example, an individual who has started to feel trapped in a relationship might in certain contexts show heightened sensitivity to feeling trapped in physically enclosed spaces based on the transfer of functions from the *relationship-trapped* relation to the *enclosed space-trapped* relation. It is difficult to see how this effect could occur based on traditional direct contingency accounts of nonhuman behavior, because the stimulus domains are formally so different, and a relevant history of differential reinforcement or stimulus pairing is completely absent. Moreover, the extended tact and the process of abstraction, that Skinner (1957) used to account for metaphor, also fails to specify how metaphor develops from this relatively simple behavioral process (see Stewart & Barnes-Holmes, 2001). In contrast, the RFT concept of relating relations provides a plausible and parsimonious account of metaphor and its role in human psychopathology.

Rule-Governed Behavior

Rules provided by others. Consider the following example of a rule provided by others. An individual planning a holiday instructs his neighbor as follows: "I'm going on vacation in two weeks and will be gone for a month. If you water and mow my lawn each week I am gone, the following month I will pay you $100" (Hayes, Gifford, & Hayes, 1998). In this case, the stated rule would appear to be simple because all of the important aspects of the contingency are stated, including: a temporal antecedent; the topography of the target response; the appropriate context for the response; the type of consequence; and when it will be delivered. Furthermore, the rule clearly alters the functions of calendar time, the grass, the lawn mower, and water. Nonetheless, the type of contingency specified here could not generate the desired behavior through direct training alone, because delayed consequences such as these are not effective without verbal rules.

For RFT, an analysis of the rule in the previous example requires: (1) an identification of the relational frames involved and the contextual cues that occasioned the relations, and (2) an analysis of the functions of the events that are transformed through these relations and the cues that occasioned the transformations. In this example, the specific relational frames involved include: coordination relations between words (e.g., lawnmower) and actual objects; before-after relations, specified in temporal terms, that function as a temporal antecedent and a consequence; and if-then relations that specify the contingent relations. In terms of the transformations of function, words such as "mow" and "water" alter the behavioral functions of the grass, and thereby provide the necessary approach functions specified in the consequence. According to RFT, therefore, rules may be defined as examples of relational networks and transformations of function that are more or less complex (Barnes-Holmes, O'Hora, Roche, Hayes,

Bisset, & Lyddy, 2001).

An individual provided with a rule can determine whether or not it is being followed by the extent to which the rule coordinates with what is being done. In more technical terms, for the listener who follows a rule, the coordination between the original relational network that constituted the rule and the relations sustained among the events specified therein, provides an on-going source of behavioral regulation. Although, rule following may involve repeating the rule, this is not essential for the rule to be followed. Because the elements specified in a rule may be actualized by the nonarbitrary environment (e.g., a sunny day one week into the neighbor's vacation), these events themselves may participate in a relational network that corresponds to the original relational network that constituted the rule. In other words, it is through these relations of coordination between relational networks that a listener is able to determine whether or not the rule is being followed. According to RFT, therefore, the term rule-governed behavior is used to describe instances in which a frame of coordination between two relational networks serves as a source of control over behavior (Barnes-Holmes, Hayes, & Dymond, 2001).

Relational Frame Theory also provides a number of reasons why rules that are stated and understood may still not be followed (Barnes-Holmes, et al., 2001). First, the absence of rule following may result from *insufficient control by nonverbal contingencies*. For example, the target behavior may not be in the behavioral repertoire of the listener, such as when a client does not possess the appropriate social skills required to perform well in job interviews.

Second, the level of rule following may depend in part on the *credibility of the speaker*. For example, a therapist with whom there is trust and intimacy may be more likely to produce effective rule following in a client than one for whom such a relationship is not present. According to RFT, this type of credibility may be acquired verbally as well as directly by experience. For example, a therapist who is regarded by a client as 'genuinely concerned for the client's well-being' may be likely to produce appropriate rule following because 'genuine concern' participates in a frame of coordination with the provision of good or accurate rules.

Third, the level of rule following may also depend on the speaker's *authority and ability to mediate reinforcement*. For example, a client is more likely to follow the rules of a therapist who discusses the rules in subsequent sessions and provides feedback or reinforcement for rule following. In contrast, a therapist who frequently provides rules without subsequently checking whether or not they have been followed may be less likely to establish appropriate rule following in clients.

Fourth, rule following may also be determined by the *plausibility of the rule* itself (see Hovland, Lumsdaine, & Sheffield, 1949). The plausibility of a particular rule may be either undermined or enhanced by the derivation of relations of distinction or opposition between the relational network constituted in the rule and other relational networks within which terms stated in the rule participate. For example, a client in therapy may be instructed to rely more heavily on her own direct experience than on the rules provided by family members. However, such a rule may be seen to challenge the cultural norm in which the family is believed to be the principal source of care and

protection. In other words, the client may derive a frame of distinction between the therapist's rule and the rule provided by the wider culture with respect to families.

Rules provided by oneself. Although the RFT interpretation of rules provided by oneself involves the same basic processes as rules provided by others, analyzing the latter also requires an understanding of the RFT concept of self. According to RFT, self-awareness involves an individual "not simply behaving with regard to his behavior, but . . also behaving verbally with regard to his behavior" (Hayes & Wilson, 1993, p. 297; see also Dymond & Barnes, 1994, 1995, 1996, 1997). Put simply, *verbally* discriminating one's own behavior that leads to self-knowledge involves relational framing, whereas simply behaving with regard to one's own behavior does not.

Relational Frame Theory argues that derived relational responding makes verbal self-knowledge important and useful (Hayes & Gifford, 1997; Hayes & Wilson, 1993). For example, humans can verbally construct a future and plan for it in great detail, thereby increasing the chances of survival. However, RFT also argues that derived relational responding makes verbal self-knowledge emotional and difficult (Hayes & Gifford, 1997; Hayes & Wilson, 1993). Consider, for example, an individual who is unexpectedly passed over for promotion. Because of the coordination relations between work and promotion, many of the functions of his daily work routine may be transformed by this failure. In this way, his promotion failure is not just experienced aversively at the time it occurs, but can be carried for days, weeks, or even months into the future. Verbally, therefore, such an individual may literally fail everyday of his working life. According to RFT, this effect would not occur without verbal relations or the transformations of function. Of course, practically all forms of psychotherapy rely upon this process. For example, simply by talking about traumatic events and thus reexperiencing the trauma, the emotional and behavioral functions of the stimuli associated with the trauma may be changed, such that one learns to be present to psychological pain rather than avoiding it. Techniques such as systematic desensitization and relaxation training are built upon this process, and focus explicitly on reducing the aversive and/or avoidance functions of the thoughts and feelings associated with trauma and its report (see Pennebaker, 1997, for a review).

Self-directed rules and the verbal construction of self. According to RFT, verbally-sophisticated individuals produce a myriad of simple and complex self-directed rules on an on-going basis. These may range from a very simple rule such as " I must make sure to get up early tomorrow to be on time for my interview" to complex rules about major life decisions or issues. The 'getting up early' rule is simply a strategic one in which the outcome is clearly and simply specified. According to RFT, if the relational terms within the network have reasonably precise behavioral functions, following the rule simply involves ensuring that the individual goes to bed at a reasonable hour, and sets the alarm clock for an appropriate time, etc. Although simple strategic problems such as this involve self-generated rules, the role of the self is somewhat limited. For example, almost anyone could instruct the listener about what to do in order to get up early tomorrow.

More extensive issues of self, however, may be involved in the overarching patterns of behavior in which the strategic rule participates. For example, the individual in question may have noticed a recent pattern of tardiness for work, for which she has no immediate explanation. Nevertheless, getting up late for work may conflict with the individual's ambitions to excel in her chosen profession. Complex issues such as these involve the self in a number of ways. First, they are all part of the on-going process of self-knowledge that includes becoming verbally aware of the time she arrives at work, feelings of guilt about being late, or ambitions about succeeding. Second, getting up on time for work may also involve a type of self-knowledge that RFT calls *the conceptualized self*, which in this case may involve a class of self-directed rules, including "I should always behave like a professional".

In contrast to strategic self-directed rules, valuative rules also play an important role in the emotional lives of human beings. Consider a woman who is questioning the future of her marriage. She may dwell upon how bad things are between her and her husband and think about leaving her husband if things do not get better between them. The relational network in this rule contains various terms that do not possess precisely controlled behavioral functions. For example, what exactly does it mean for 'things to get better'? Furthermore, the outcomes of the rule following itself cannot be known. For example, what exactly will be the outcomes of either staying in or leaving the marriage?

At least two generic strategies are available to the individual in the context of self-directed rules that are primarily valuative. One would involve the woman engaging in additional relational activities that would bring her into contact psychologically and emotionally with a range of possible outcomes. For instance, a therapist might ask the woman to try to imagine what it would be like on a day-to-day basis living alone without her husband. The other strategy would involve making psychological and emotional contact with different features of the current problem. For example, a therapist might ask the woman to imagine what her marriage would be if it was an animal, and to explore the various properties of the animal and what they might represent for her.

According to RFT, self-directed rules that involve valuative problem solving require a highly complex sense of self that can be understood in terms of what RFT calls *the three selves*. Relational Frame Theory defines these selves as: the *conceptualized self* (i.e., the self as the content of verbal relations); the *knowing self* (i.e., the self as the ongoing process of verbal relations); and the *transcendent self* (i.e., the self as the context of verbal relations, see Hayes, 1995 and Barnes-Holmes, et al., 2001). If the woman chooses to end the marriage largely on the basis of self as content, little or no contact may be made verbally with important consequences of her actions. For example, she may decide to stay in the marriage based on the self-directed rule that she is the type of woman who simply cannot live without a man in her life. In this case, the verbal construction of self (as content) as "needing a man" dominates her problem solving, rather than the verbally constructed future of a life of misery in a failed marriage.

A decision to act based solely on self as process may also be problematic. For example, the woman may decide to leave the marriage prematurely because she has accurately recognized that she has been unhappy for some time. However, such a

decision, based as it is entirely on self as process, does not involve constructing relational networks in which self as content plays a role. Thus, the woman may leave the marriage for purely emotional reasons without asking herself if she wants to be the type of person who leaves a marriage, for example, 'when the going gets tough'?

Alternatively, the woman may act on the basis of self as context, thereby providing a stable perspective from which a more balanced decision can be made. Self as context provides a psychological space in which she can still contact both self as content (e.g., I am the type of woman who needs a man), and self as process (e.g., but I have feelings of hostility towards my husband all the time), but neither obtain absolute control over her decision. In this way, a broader range of relevant issues are contacted and dealt with in the final choice. For example, in contacting both the need to be married and her feelings of hostility, this may help the woman to see that the former is feeding the latter. In other words, in an effort to protect her self as content (i.e., to stay married), the woman has avoided expressing honest feelings of frustration with her husband out of fear that he might leave her, and as a result, these feelings of frustration have grown into intense hostility. It is only from self as context, that this dynamic between her other two selves can be discriminated, and thus form a useful basis for future action.

PART 3

A RELATIONAL FRAME THEORY INTERPRETATION OF ACCEPTANCE, DEFUSION, AND VALUES

Having outlined the basic analytic concepts of RFT and how they can be used to interpret certain psychological problems, we will now offer a preliminary and brief interpretation of three core features of Acceptance and Commitment Therapy (ACT - see Hayes, et al. 1996). We focus on this therapy, in particular, because it articulates directly with RFT, and we are each concerned with pursuing research programs that are designed to lay out a more complete analysis of the relationship between RFT and ACT. Indeed, we see the two as inextricably intertwined and each is necessary for the other's ongoing development. In what follows, however, it is important to recognize that ACT is a complex and multi-faceted form of psychotherapy, and thus the current interpretations must be seen as no more than a heuristic for further conceptual and empirical RFT analyses of ACT (e.g., Pankey & Hayes, 2003).

Acceptance

Acceptance is not a technical term, but is useful in that it serves to orient us towards an important domain in the understanding and treatment of human suffering. From an RFT perspective, when acceptance is stripped down to its core verbal relational elements, it seems to involve the following: (1) a rule or relational network in which a possibly aversive event is made contingent on a particular behavior (e.g., "if I go to the mall, I will feel panicky"); (2) an action that then brings the individual into contact with that aversive event (i.e., going to the mall); (3) the construction of a relational network that coordinates with that contact (e.g., "I am at the mall and I feel panicky);

and (4) a frame of coordination occurs between this network and the original rule or relational network (i.e., "I was right, I said that I would feel panicky, and I did"). From an RFT point of view, the coordination between relational networks that occurs under point 4 above plays an important role in the apparent efficacy of acceptance-based therapeutic strategies. Because the two networks cohere, and relational coherence has a high probability of reinforcement within the verbal community, the likelihood of following an acceptance-based strategy in the future is increased. For example, the panic disordered client may be more likely to continue with a shopping trip if he is told that he may well experience feelings of intense panic and still continue shopping, rather than being told that his fears are irrational and that nothing bad can happen.

As another example, consider the following experimental analog based on a recent study (Gutiérrez, Luciano, Rodríguez, & Fink, in press). A female participant is told that she is going to receive electric shocks of increasing intensity, and that she will perceive them to be very painful, but despite this pain, she can continue with the experimental task. In this case, the original rule –"this is going to hurt"– may coordinate with her own verbal construction of what happens to her in the experiment, and this coherence between the two networks makes it more likely that she will respond in accordance with the additional part of the rule –"although it hurts, you can continue with the task". In contrast, imagine that the woman was told that if she engaged in some form of thought suppression, she could control the level of pain. In this case, when the shocks become quite intense, a frame of distinction may emerge between the rule provided by the experimenter and the relational network generated by the participant. At this point, the relational incoherence created between the experimenter's and the participant's rules may lead the woman to end the experimental task. More informally, the woman might think: "You were wrong, I can't control the pain and I want to stop now". This thought experiment provides a reasonable analog of the way in which acceptance strategies may be used to cope with human psychopathology.

Defusion

Of course, as any therapist knows, successful therapy requires a great deal more than simply telling clients that they are sometimes going to feel very bad, but they should get on with their lives anyway. For this basic message to work, the ACT therapist must focus considerable energy on assisting the client to defuse from their thoughts and feelings in the service of achieving valued objectives. A range of exercises and metaphors are typically used in ACT to facilitate defusion. For example, the use of the verbal convention "I am having the thought that..." may be employed to highlight the nonliteral quality of the client's thoughts. Exercises such as 'The Floating Leaf,' in which clients are asked to place every thought or feeling that comes up on a leaf and watch it float down a stream, are also used. The purpose of these and many other exercises and metaphors, from an RFT perspective, is to transform the relational functions of thoughts and feelings from I HERE NOW to I THERE THEN. When we respond to our thoughts as I HERE NOW, we are in a sense caught up in them –we are not perceiving those thoughts and feelings *from* a perspective, but rather they *are* the perspective. However,

when we respond to our thoughts and feelings as I THERE THEN, they are seen as thoughts and feelings that we had, rather than providing our current and only perspective. Thus, for example, a client may be encouraged to treat her thoughts and feelings as *just* thoughts and feelings (i.e., "I just had the thought that I am a bad person" rather than "I am a bad person"). From the RFT point of view, an important part of therapy, including homework, involves extensive exemplar training in the appropriate transformations of relational functions, such that the typical avoidance controlling properties of negative thoughts and feelings are somewhat reduced in order to facilitate client participation in acceptance-based strategies.

Values

For ACT therapists, acceptance-based strategies and defusion techniques are not ends in themselves –these two therapeutic elements are synergized in the service of living a valued life. An important part of ACT, therefore, involves helping clients to make contact with their own values, so that they can focus acceptance and defusion strategies on working towards those values. From an RFT point of view, a life value may be interpreted as the highest point in a hierarchical relational network. For example, valuing family life is the type of value that is not typically subsumed under a higher point in the network. One values family life simply because one values family life -no further justification is required, and in this sense we have reached the top of the hierarchy. In contrast, earning money typically participates at a lower point in a values hierarchy because it serves values that participate at higher points in the hierarchy, such as family life, holidays, and hobbies, etc.

An important part of ACT involves helping clients to construct value hierarchies that reflect their histories and current contexts. So, for example, a disillusioned surgeon may, during the course of therapy, come to realize that she pursued a medical career simply because she wanted to please her father. In this case, she may have failed to contact at least two of her values: (1) pursuing her own chosen career, irrespective of her father's perceived wishes, and (2) the possibility of a more intimate relationship with her father in which she can be honest about what she really wants to do with her life. And, of course, making contact with these values and acting in accordance with them, will likely require both acceptance and defusion as defined above. Telling her father that she wishes to leave medicine is likely to make the client feel very uncomfortable. In effect, a therapeutic rule or relational network is constructed that predicts an aversive psychological event, and indeed the client will likely experience this when she faces her father. In order to accept the discomfort that shows up in this situation, exemplar training in defusion will serve to undermine the avoidance responses that have led her away from her chosen values. In this way, the dynamic interplay among acceptance, defusion, and values provides a focus for the ACT and RFT approaches to human psychopathology.

CONCLUSION

In the current paper, we have attempted to show how both the basic and applied sciences of behavior analysis have been transformed by the modern research agenda in human language and cognition, focusing in particular on RFT. The traditional behavioral view of human psychological problems, in terms of direct operant and respondent contingency analyses, appears inadequate in a modern light. Indeed, as we have aimed to show, a more complete analysis and understanding of human psychopathology requires that we deal with relational frames, relational networks, relating relations, rules, perspective-taking, and the concept of self. The applied science of behavior therapy will necessarily be transformed by taking on board this new and revolutionary way of looking at human suffering (see also Harrington, Fink, & Dougher, 2001). But we are only beginning. Much of the basic studies in RFT are focused on very simple relational networks, and only the most recent studies have examined metaphor, rule-governance, and perspective-taking. Moreover, fresh new research programs are needed that directly target RFT analyses of the key psychological processes involved in ACT. These are indeed exciting times.

REFERENCES

American Psychiatric Association (1994). *Diagnostic and statistical manual of mental disorders*, (4th edition). Washington, DC: American Psychiatric Association.

Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Review, 84*, 191-215.

Bandura, A. & Menlove, F.L. (1968). Factors determining vicarious extinction of avoidance behavior through symbolic modeling. *Journal of Personality and Social Psychology, 8*, 99-108.

Barlow, D.H., Hayes, S.C., & Nelson, R.O. (1985). *The scientist practitioner: Research and accountability in clinical and educational settings*. Oxford: Pergamon.

Barnes, D., Browne, M., Smeets, P.M., & Roche, B. (1995). A transfer of functions and a conditional transfer of functions through equivalence relations in three- to six-year old children. *The Psychological Record, 45*, 405-430.

Barnes, D., Lalor, H., Smeets, P., & Roche, B. (1996). Stimulus equivalence and academic self-concept among mildly mentally handicapped and nonhandicapped children. *The Psychological Record, 46*, 87-107.

Barnes, D. & Roche, B. (1996). Relational frame theory and stimulus equivalence are fundamentally different: A reply to Saunders. *The Psychological Record, 46*, 489-508.

Barnes, D. & Roche, B. (1997b). Relational frame theory and the experimental analysis of human sexuality. *Applied and Preventive Psychology, 6*, 117-135.

Barnes-Holmes, D. & Barnes-Holmes, Y. (2000). Explaining complex behavior: Two perspectives on the concept of generalized operant classes. *The Psychological Record, 50*, 251-265.

Barnes-Holmes, Y., Barnes-Holmes, D., Roche, B., Healy, O., Lyddy, F., Cullinan, V., & Hayes, S.C. (2001). In. S.C. Hayes, D. Barnes-Holmes & B. Roche (Eds.), *Relational frame theory: A post-Skinnerian account of language and cognition*. New York: Plenum Press.

Barnes-Holmes, Y., Barnes-Holmes, D., Roche, B., & Smeets, P.M. (2001a). Exemplar training and a

derived transformation of function in accordance with symmetry. *The Psychological Record, 51*, 287-308.

Barnes-Holmes, Y., Barnes-Holmes, D., Roche, B., & Smeets, P.M. (2001b). Exemplar training and a derived transformation of function in accordance with symmetry: II. *The Psychological Record, 51*, 589-604.

Barnes-Holmes, D., Dymond, S., Roche, B., & Grey, I. (1999). Language and cognition. *The Psychologist, 12*, 500-504.

Barnes-Holmes, D., O'Hora, D., Roche, B., Hayes, S.C., Bisset, 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 language and cognition* (pp. 103-118). New York: Plenum.

Barnes-Holmes, D., Hayes, S.C., & Dymond, S. (2001). Self and self-directed rules. In. S.C. Hayes, D. Barnes-Holmes & B. Roche (Eds.), *Relational frame theory: A post-Skinnerian account of language and cognition* (pp. 119-139). New York: Plenum Press.

Catania, A.C. (1998). *Learning* (4th Edition). Upper Saddle River, NJ: Prentice-Hall.

Chomsky, N. (1959). A review of B.F. Skinner's Verbal Behavior. *Language, 35*, 26-58.

Davison, G.C., Neale, J.M., & Krung, A.M. (2004). *Abnormal psychology* (9th ed.). New Jersey: Wiley.

deRose, J.T., deSouza, D.G., Rossito, A.L., & deRose, T.M.S. (1992). Stimulus equivalence and a generalization in reading after matching to sample by exclusion. In S.C. Hayes & L.J. Hayes (Eds.), *Understanding verbal relations* (pp. 69-82). Reno NV: Contest Press.

Dymond, S. & Barnes, D. (1994). A transfer of self-discrimination response functions through equivalence relations. *Journal of the Experimental Analysis of Behavior, 62*, 251-267.

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 the Experimental Analysis of Behavior, 64*, 163-184.

Dymond, S. & Barnes, D. (1996). A transformation of self-discrimination response functions in accordance with the arbitrarily applicable relations of sameness and opposition. *The Psychological Record, 46*, 271-300.

Dymond, S. & Barnes, D. (1997). Behavior analytic approaches to self-awareness. *The Psychological Record, 47*, 181-200.

Ellis, A. (1958). Rational psychotherapy. *Journal of General Psychology, 59*, 35-49.

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*, 137-156.

García, A. & Benjumea, S. (2001). Prerrequisitos ontogenéticos para la emergencia de relaciones simétricas. *International Journal of Psychology and Psychological Therapy, 1,* 115-135.

Gutiérrez, O., Luciano, C., Rodríguez, M., & Fink, B.C. (in press). Comparison between an acceptance-based and a cognitive-control-based protocol for coping with pain. *Behavior Therapy*.

Hall, J.F. (1984). Backward conditioning in Pavlovian type studies: Reevaluation and present status. *Pavlovian Journal of Biological Sciences, 19*, 163-168.

Harrington, A., Fink, B., & Dougher, M. (2001). Into the lion's den: Incorporating personality and evolutionary psychology to expand clinical behavior analysis. *International Journal of Psychology and Psychological Therapy, 1,* 175-189.

Hayes, S.C. (1989) (Ed). *Rule-governed behavior: Cognition, contingencies, and instructional control*. New York: Plenum.

Hayes, S.C. (1995). Knowing selves. *The Behavior Therapist, 18*, 94-96.

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

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

Hayes, S.C. & Gifford, E.V. (1997). The trouble with language: Experiential avoidance, rules, and the nature of verbal events. *Psychological Science, 8*, 170-173.

Hayes, S.C., Gifford, E.V., & Hayes, G.J. (1998). Moral behavior and the development of verbal regulation. *The Behavior Analyst, 21*, 253-279.

Hayes S.C., Gifford, E.V., & Ruckstuhl, Jr., L.E. (1996). Relational frame theory and executive function. In G.R. Lyon & N.A. Krasnegor, (Ed's.), Attention, memory and executive function (pp.279-305). Baltimore: Brookes.

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). New York: Plenum.

Hayes, S.C., Nelson, R.O., & Jarrett, R. (1987). Treatment utility of assessment: A functional approach to evaluating the quality of assessment. *American Psychologist, 42*, 963-974.

Hayes, S.C., Strosahl, K., & Wilson, K.G. (1999). Acceptance and Commitment Therapy: An experiential approach to behavior change. Guilford Press, New York.

Hayes, S.C. & Toarmino, D. (1999). The Rise of Clinical Behaviour Analysis. *The Psychologist, 12*, 505-509.

Hayes, S.C. & Wilson, K.G. (1993). Some applied implications of a contemporary behavior-analytic account of verbal events. *The Behavior Analyst, 16*, 283-301.

Hayes, S.C., Wilson, K.G. Gifford, E.V., Follette, V.M., & Strosahl, K. (1996). Experiential avoidance and behavioral disorders: A functional dimensional approach to diagnosis and treatment. *Journal of Consulting and Clinical Psychology, 64*, 1152-1168.

Hovland, C.I., Lumsdaine, A.A., & Sheffield, F.D. (1949). *Experiments on mass communication: Studies in social psychology in World War II*, Vol. 3. Princeton, NJ: Princeton University Press.

Leader, G., Barnes, D., & Smeets, P.M. (1996). Establishing equivalence relations using a respondent-type training procedure. *The Psychological Record, 46*, 685-706.

Lovaas, O.I. (1981). *Teaching developmentally disabled children: The ME book*. Texas: Pro-ed.

Mahoney, M.J. (1994). Cognition and behavior modification. Cambridge, MA: Ballinger.

Mahoney, M.J. & Arnkoff, D.B. (1978). Cognitive and self-control therapies. In S. L. Garfield & A. E. Bergin, (Eds.), *Handbook of psychotherapy and behavior change* (2nd edition). N. York: Wiley.

Malan, D.H. (1995). *Individual psychotherapy and the science of psychodynamics*. Oxford: Butterworth Heinemann.

Moorey, S. (2003). Cognitive therapy: Historical context and development in Britain. In W. Dryden (Ed.), *Handbook of individual therapy* (fourth edition) (pp. 294-323). London: Sage.

Pankey, J. & Hayes, S.C. (2003). Acceptance and Commitment Therapy for psychosis. *International Journal of Psychology and Psychological Therapy, 3,* 311-328.

Pennebaker, J.W. (1997). Writing about emotional experiences as a therapeutic process. *Psychological Science, 8*, 162-166.

Reese, H.W. (1968). *The perception of stimulus relations: Discrimination learning and transposition*.

New York: Academic Press.

Roche, B., & Barnes, D. (1997). A transformation of respondently conditioned functions in accordance with arbitrarily applicable relations. *Journal of the Experimental Analysis of Behavior, 67*, 275-301.

Roche, B. & Barnes, D. (1998). The experimental analysis of human sexual arousal: Some recent developments. *The Behavior Analyst, 21*, 37-52.

Roche, B., Barnes-Holmes, D., Smeets, P.M., Barnes-Holmes, Y., & McGeady, S. (2000). Contextual control over the derived transformation of discriminative and sexual arousal functions. *The Psychological Record, 50*, 267-291.

Seligman, M.E.P. (1974). Depression and learned helplessness. In R.J. Friedman, & M.M. Katz, (Eds.), *The psychology of depression: Contemporary theory and research*. Washington, DC: Winston-Wiley.

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

Sidman, M. (1994). *Stimulus equivalence: A research story*. Boston: Authors Cooperative.

Skinner, B.F. (1938). *Behavior of organisms*. Appelton-Century-Crofts, New York.

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

Smeets, P.M., Barnes, D., & Roche, B. (1997). Functional equivalence in children. Derived stimulus-response and stimulus-stimulus relations. *Journal of Experimental Child Psychology, 66*, 1-17.

Stewart, I. & Barnes-Holmes, D. (2001). Understanding metaphor: A relational frame perspective. *The Behavior Analyst, 24*, 191-199.

Stewart, I., Barnes-Holmes, D., Hayes, S.C., & Lipkens, R. (2001). Relations among relations: Analogies, metaphors and stories. In. S. C. Hayes, D. Barnes-Holmes, & B. Roche (Eds.), *Relational frame theory: A post-Skinnerian account of language and cognition*. New York: Plenum Press.

Watson, J.B. & Rayner, R. (1920). Conditioned emotional reactions. *Journal of Experimental Psychology, 3*, 1-14.

Wilson, G.T. & Davidson, G.C. (1971). Processes of fear reduction in systematic desensitisation: Animal studies. *Psychological Bulletin, 76*, 1- 14.

Wilson, K.G., Hayes, S.C., Gregg, J., & Zettle, R.D. (2001). Psychopathology and psychotherapy. In: S.C. Hayes, D. Barnes-Holmes & B. Roche (Eds.), *Relational frame theory: A post-Skinnerian account of language and cognition*. New York: Plenum Press.

Wolpe, J. (1952). Objective psychotherapy of the neuroses. *South African Medical Journal, 26*, 825-829.

Wolpe, J. (1954). Reciprocal inhibition as the main basis for psychotherapeutic effects. *Archives of Neurological Psychiatry, 72*, 205–226.

Wolpe, J. (1958). Psychotherapy by reciprocal inhibition. Standford, CA: Stanford University Press.

Wulfert, E. & Hayes, S.C. (1988). The transfer of conditional sequencing through conditional equivalence classes. *Journal of the Experimental Analysis of Behavior, 50*, 125-144.

Visdómine, J.C. & Luciano, M.C. (2002). Formación de locus de control por transferencia de funciones. *International Journal of Psychology and Psychological Therapy, 2,* 57-73.