A Contextual Behavioral Approach to the Study of (Persecutory) Delusions

Corinna Stewart and Ian Stewart

National University of Ireland, Galway

Sean Hughes

Ghent University

Author Note

CS and IS, Department of Psychology, National University of Ireland, Galway, Ireland. SH, Department of Experimental Clinical and Health Psychology, Ghent University, Belgium. Preparation of this paper was supported by a NUI Galway Doctoral Scholarship to CS. Electronic mail should be sent to Corinna Stewart at c.stewart2@nuigalway.ie

Abstract

Throughout the past century the topic of delusions has mainly been studied by researchers operating at the mental level of analysis. According to this perspective, delusional beliefs, as well as their emergence and persistence, stem from an interplay between (dysfunctional) mental representations and processes. Our paper aims to provide a starting point for researchers and clinicians interested in examining the topic of delusions from a functional-analytic perspective. We begin with a brief review of the research literature with a particular focus on persecutory delusions. Thereafter we introduce Contextual Behavioral Science (CBS), Relational Frame Theory (RFT) and a behavioral phenomenon known as arbitrarily applicable relational responding (AARR). Drawing upon AARR, and recent empirical developments within CBS, we argue that (persecutory) delusions may be conceptualized, studied and influenced using a functional-analytic approach. We consider future directions for research in this area as well as clinical interventions aimed at influencing delusions and their expression.

Keywords: Persecutory Delusions, Contextual Behavioral Science, RFT

A Contextual Behavioral Approach to the Study of (Persecutory) Delusions

While of much interest in clinical, cognitive, and neuropsychology, delusions have yet to be systematically explored within Contextual Behavioral Science (CBS). In Part I we introduce the mainstream literature on delusions and its focus on the identification of mediating mental mechanisms. In Part II we examine how delusions may be approached from a CBS (functional-analytic) perspective. This latter approach sets the idea of mental constructs as causal events to the side. Instead it defines delusional beliefs as behaviors and investigates what environmental factors might influence them. We then introduce Relational Frame Theory (RFT; Hayes, Barnes-Holmes, & Roche, 2001), which views arbitrarily applicable relational responding (AARR) as the key functional unit underlying human language and cognition. In Part III we forward the idea that delusions can be conceptualized in purely functional terms using AARR and related concepts. Finally, in Part IV, we consider future research directions, challenges, and clinical implications of a functional approach to delusions.

Note that this paper is primarily a "call to arms" for a systematic, inductive, and empirically informed functional analysis of "delusional beliefs". Our aim is not to present a theoretical account of delusions nor to review decades of functional research on this topic. In fact, we highlight the absence of any such theorizing or research within CBS. We instead want to draw attention to delusions as a topic for study and to provide a possible starting point for CBS researchers and clinicians interested in studying them. We present possible directions for research and key questions that may shape a functional approach to (persecutory) delusions in the years to come.

Part I: A Brief Introduction to the Study of Delusions

Definitional Issues

Previous research tended to examine delusions and related phenomena under the rubric of psychotic disorders, including schizophrenia, schizoaffective disorder, and delusional disorder. However, "these diagnoses may prove an obstacle in the advancement of the understanding and treatment of the difficult experiences for which patients require help. The empirical research indicates that within these diagnoses are multiple independent experiences" (Freeman & Garety, 2014, p. 1180). Consequently, researchers increasingly believe that focusing on specific experiences (e.g., delusions, hallucinations), rather than diagnostic classifications (e.g., schizophrenia), may yield more precise accounts of the etiology of, and processes implicated in, these phenomena (e.g., Beards & Fisher, 2014; Bentall, 2006; Freeman, Garety, Kuipers, Fowler, & Bebbington, 2002).

The definition of what constitutes a 'delusion' has shifted dramatically across time, tightly linked to the syndromal approach adopted in the *Diagnostic and Statistical Manual of Mental Disorders* (*DSM*). In the most recent version of the *DSM* (*DSM-V*), delusional beliefs are (syndromally) defined as distorted or excessive inferences about reality highly resistant to change despite conflicting evidence (APA, 2013). Yet recent research indicates that delusional beliefs constitute a more complex phenomenon[1]. Freeman (2007) outlined a range of factors relevant when determining whether a delusional belief is present, including the extent to which the belief (a) involves personal reference, (b) is unfounded (though possibly reflecting a kernel of truth that has been exaggerated or distorted), (c) is firmly held and resistant to change (even given contradictory evidence), (d) is preoccupying and/or distressing, and (e) impairs functioning or produces other negative life-outcomes (e.g., distress, aggression) (also see Oltmanns, 1988).

Persecutory Delusions

[1] For instance, recent work suggests that the dimensions of distress, preoccupation, and conviction, as well as appraisals of the experiences and coping strategies adopted, may be more relevant than the content of delusional beliefs alone (e.g., Peters, Joseph, & Garety, 1999; Peters, Joseph, Day, & Garety, 2004).

There are several sub-types of delusions such as erotomanic (e.g., believing that a stranger, usually someone famous, loves me), grandiose (e.g., believing I have superior qualities or abilities), jealous (e.g., wrongly believing my partner is unfaithful), somatic (e.g., believing I have a physical defect or medical problem), mixed-type (e.g., having more than one type of delusion), and unspecified. The most prevalent subtype, however, is persecutory delusions (PD), wherein a person believes they are "being conspired against, cheated, spied on, followed, poisoned or drugged, maliciously maligned, harassed, or obstructed in the pursuit of long-term goals" (APA, 2013, p. 90). Freeman and Garety (2000) suggest that PD beliefs have two core defining properties: the individual believes (a) harm is occurring or will occur, and (b) a perceived persecutor is deliberately trying to cause that harm. Interestingly, evidence suggests such 'paranoid' thoughts and ideas are not exclusive to those with severe mental illness. Rather there seems to be an exponential distribution of paranoid thoughts within the general population (Bebbington et al., 2013; Freeman et al., 2005), with paranoid thoughts and ideas occurring along a continuum, with normal experiences at one end and PD at the other, such that "many people have a few paranoid thoughts, and a few have many" (Freeman & Garety, 2014, p.1179).

The functional perspective outlined later in this paper applies to the study of delusions in general but we decided to focus on PD for several reasons. First, PD are the most prevalent subtype of delusional beliefs (APA, 2000) such that 50 to 70 per cent of patients present with PD at first episode of psychosis (Freeman & Garety, 2014; Sartorius, et al., 1986). Second, PD are the most common delusion-type across cultures (Moutoussis, Williams, Dayan, & Bentall, 2007) and present in many different disorders (e.g., schizophrenia, bipolar, PTSD, dementia). Third, PD are related to a wide assortment of clinical phenomena including anxiety, depression, poor health, suicidal ideation, poor social functioning, substance abuse, as well as increased use of mental health services (Birchwood, Iqbal, Chadwick, & Trower,

2000; Freeman et al., 2011). Indeed, almost half those with PD have levels of psychological wellbeing in the lowest 2% of the general population (Freeman et al., 2014). Finally, PD are the delusion subtype most likely to be acted upon (Freeman et al., 2007).

Origins of PD

A number of genetic (and epigenetic), neurological (e.g., neurotransmitter functioning), social (e.g., interpersonal trauma), cognitive (e.g., impaired Theory of Mind), and environmental (e.g., prenatal stress, urbanization) factors have been implicated in the etiology of PD. For instance, Collip, Myin-Germeys, and Van Os (2008) argue that exposure to particular environmental events, in interaction with (epi)genetic factors, may induce psychological or physiological alterations that can be traced to a final common pathway of cognitive biases and/or altered dopamine neurotransmission ("sensitization"), which may facilitate the onset and persistence of psychotic symptoms, including PD (also see Read, Fosse, Moskowitz, & Perry, 2014). While theories concerning physiological, genetic, and neurological factors in PD are important (Freeman, Bentall & Garety, 2008) they are beyond the remit of this paper (which focuses on psychology).

In what follows we consider empirical and theoretical trends currently shaping the study of PD in psychological science. We first discuss three cognitive models concerned with the origins and persistence of PD including their strengths and limitations. We then turn to the functional-analytic approach.

Impaired Theory of Mind. Theory of Mind (ToM) is the ability to infer the mental states of others and anticipate their beliefs, intentions, and behavior. Several authors have argued that ToM abilities are impaired in individuals with schizophrenia in general and PD in particular. Frith (1992) argued that impaired ToM may be linked to three major groups of schizophrenic features (symptoms). The first group ('disorders of willed action') refers to inability to perceive one's intentions or consider one's behavior as the product of one's

actions. The second group ('impaired self-monitoring') refers to inability to recognize situations in which one's thoughts may lead to delusional beliefs or hallucinations. The third group ('impaired monitoring of others') refers to a lack of awareness or impaired understanding of the thoughts or intentions of others. The individual can still infer other people's thoughts or intentions but these inferences are typically incorrect. Frith argued that this third type of ToM impairment is linked to PD.

Several studies provide evidence that individuals suffering from paranoid schizophrenia and PD also suffer from impaired ToM (e.g., Corcoran, Mercer, & Frith, 1995; Frith & Corcoran, 1996; Mehl et al., 2010; Scherzer, Leveille, Achim, Boisseau, & Stip, 2012). However, others have found that those reporting delusional beliefs score similarly on ToM tasks to those who do not (Drury, Birchwood, & Robinson, 1998; Greig, Bryson, & Bell, 2004; see also Korver-Nieberg et al., 2013). Indeed, following a review of 38 studies, Garety and Freeman (2013) stated that the majority did not find any association between ToM and delusions in general or PD in particular. Thus, while some individuals reporting PD may display ToM impairments, little evidence exists linking impaired 'monitoring of others' to PD. Indeed, Salvatore and colleagues (2012) recently argued that impairments in the ability to infer the intentions of others does not explain why those with PD tend to view those intentions as malevolent rather than neutral or positive. They suggest ToM dysfunction may contribute to inability to reflexively re-evaluate firmly held beliefs (Drury et al., 1998; Garety & Freeman, 1999).

Threat anticipation model. Freeman et al.'s (2002) multifactorial model of PD has been very influential in the field. It argues PD arises from a complex interaction between internal physiological experiences (e.g., heightened arousal, sleep disturbance, illicit drug use), anomalous experiences (e.g., perceptual anomalies, depersonalization), external experiences (e.g., verbal and non-verbal behavior of others), and cognitive reasoning biases

(e.g., jumping to conclusions). PD is argued to arise from the search for meaning following anomalous experiences (e.g., unexplained anxious arousal, depersonalization, perceptual disturbances, etc.). These experiences may be triggered *directly* by stressful events (e.g., difficult interpersonal relationships, isolation), *indirectly* by emotional disturbances (e.g., anxiety), or as a result of cognitive and reasoning biases (e.g., conformation bias). At the same time, pre-existing schemas or beliefs concerning the self, others, and the world are also incorporated into this search for meaning. For instance, a persecutory belief is likely to be formed if the individual already believes they are vulnerable, or because they view the world as hostile and threatening, perhaps based on previous traumatic experiences (see Beard & Fisher, 2014; Bentall et al., 2014). Freeman and colleagues propose that a heightened emotional state may lead the individual to "incorporate internal and external events that are unusual, ambiguous, negative, or neutral (and often socially significant) into this search for an explanation of the anomalous experiences" (p. 334). For example, they may interpret their feelings (e.g., anxiety) together with external experiences (e.g., facial expressions of strangers) and cognitions (e.g., "others don't like me") as evidence of threat. They also argue that a delusion may reflect an attribution which in turn leads to further attributions. For example, someone may attribute another's glance at them as "they're watching me", and this may facilitate subsequent delusional beliefs (e.g., "some people are out to get me").

Defensive model of PD. Bentall and colleagues (1994; 2001) argue that PD emerge as a result of defensive efforts against negative affective processes and cognitions rather than as a direct reflection of emotional concerns. Bentall, Kinderman, and Kaney (1994) proposed that in order to preserve self-esteem, and avoid discrepancies between actual and ideal self, individuals with PD avoid activation of negative self-beliefs by making externalizing, personalizing attributions for negative events (i.e., attributional biases). Bentall et al. (1994) also argued that PD sufferers have latent negative beliefs about the self and reasoned that

studies showing a discrepancy between implicit and explicit self-esteem might support their model (e.g., Valiente et al., 2011).

In their review, Garety and Freeman (2013) found that individuals with PD do demonstrate attributional biases; they show an externalizing bias for negative events when confronted with self-related (rather than neutral or abstract) stimuli and are more likely to blame people rather than situations or chance for bad events. While some studies suggest those with PD display an exaggerated self-serving bias – attributing positive events excessively to internal causes and negative events excessively to external causes (e.g., Kaney & Bentall, 1989; 1992) – others have replicated only the latter pattern (e.g., Lyon, Kaney, & Bentall, 1994). The literature on discrepancies between implicit and explicit self-esteem is also mixed. In a recent meta-analysis, Kesting and Lincoln (2013) concluded that both patients with PD and individuals with delusion-proneness in the general population have lower global self-esteem. Moreover, results from studies investigating explicit self-esteem, implicit self-esteem, and their discrepancy do not suggest that PD enhance self-esteem. Freeman and Garety (2014) highlight that a number of systematic reviews indicate that PD is directly associated with negative self-concepts, "without the need to evoke defensive processes" (p. 1181).

Persistence of PD

Cognitive theories not only tackle the origins of PD but also factors responsible for their persistence across time and context. One such factor is cognitive or reasoning biases (Freeman, 2007). Paranoid thoughts and feelings are evident in both clinical and non-clinical populations and severity varies along a continuum. For those high on this continuum, otherwise normal cognitive processes (e.g., reasoning) can break down in certain contexts or with specific stimuli.

Consider the 'jumping to conclusions' (JTC) bias which refers to situations where people reach a decision or accept an explanation despite limited information (e.g., Garety et al., 2011; So et al., 2012). Individuals with delusions are more likely to show evidence of JTC (e.g., Fine, Gardner, Craigie, & Gold, 2007) and are often more confident in their incorrect decisions than others (e.g., Moritz & Woodward, 2006; Moritz, Woodward, & Rodriguez-Raecke, 2006). Several researchers argue that the tendency to limit data-gathering before reaching a conclusion increases acceptance of delusional ideas, and may also help maintain such beliefs (Freeman & Garety, 2014).

Confirmatory and disconfirmatory biases may also feed into PD. Most people engage in confirmatory biases occasionally, including those with PD. Yet given that the latter are highly focused on detecting danger, confirmatory biases further increase the likelihood that (ambiguous) events are seen as threatening. This probability may be further compounded by memory biases (which lead to frequent recall of confirmatory evidence for the delusion; Freeman et al., 2002), a bias against disconfirmatory evidence (Moritz & Woodward, 2006), as well as a reduced tendency to adopt experiential and rational reasoning styles (Freeman, Lister, & Evans, 2014). Social factors are also important: if interactions with others are incorporated into the delusional belief (e.g., treating others with suspicion, becoming hostile), then negative reactions from others might also confirm it. Additionally, the tendency to engage in safety behaviors such as avoidance, escape, or aggression, may simultaneously prevent people from encountering disconfirmatory evidence. When someone engages in safety behaviors to avoid or escape a feared outcome, and that outcome does not occur, they may (incorrectly) conclude the safety behaviors prevented harm, and that the situation was a "near miss" (Freeman et al., 2002, p. 339). This coheres with the finding that PD sufferers had emitted at least one safety behavior in the previous month (typically avoidance), and that anxiety was correlated with safety behaviors (Freeman et al., 2013).

Persistence of delusional beliefs has also been linked to 'belief inflexibility' (inability to evaluate a belief as incorrect and consider alternatives; Freeman & Garety, 2014; Garety et al., 1997; So et al., 2012). However, while delusions are generally considered 'fixed' or resistant to change, this is not always the case. Garety et al. (2005) found roughly half of patients with delusions acknowledged they might be mistaken while Freeman et al. (2004) found 25% of patients with delusions could generate alternative explanations for their anomalous experiences. Those who could not do so reported significantly more internal anomalous experiences and displayed a more 'hasty' reasoning style. In contrast, those who reported they may be mistaken showed improvements on symptom severity (Garety et al., 1997). Thus, allowing that one's (delusional) belief might be incorrect, confronting disconfirmatory evidence, and generating alternative explanations all appear important in belief flexibility (e.g., Freeman et al., 2004; Garety et al., 1997).

Summary

Cognitive accounts have attributed the origins of delusions to a host of mental processes: from impaired ToM abilities and defensive efforts against negative affect and self-judgments, to attempts to make sense of anomalous experiences. These models also suggest that delusions persist due to cognitive and reasoning biases (e.g., JTC, belief inflexibility, confirmatory bias) and specific behavioral responses (e.g., safety behaviors). Such theories have rapidly accelerated our basic understanding of PD and provide invaluable information about the experiences (e.g., childhood, anomalous, and interpersonal experiences) that contribute to development and persistence of delusions.

Nevertheless, this work has largely been conducted by researchers operating at the mental level of analysis, for whom "psychological events are…similar to a machine, composed of discrete parts that interact and are subject to specific operating conditions. The primary scientific goal involves identifying the mental mechanism(s) that mediate between

input (environment) and output (behavior). The researcher's role is to develop an account of mental representations and processes that mediate changes in behavior" (Hughes, De Houwer, & Barnes-Holmes, 2016, p.23).

Such an analytic strategy can identify relationships between different psychological phenomena that can be used to predict future behavior. Yet it can orient attention away from those environmental variables and histories of learning that – in principle – allow the researcher to predict-and-influence the phenomenon of interest. For example, Garety, Hemsley, and Wessely (1991) found that individuals with delusional beliefs tend to exhibit evidence of JTC on probabilistic reasoning paradigms. Although a correlation between cognitive reasoning biases and PD improves our understanding of PD, it does not bring us closer to understanding the environmental variables or history of learning that give rise to JTC or PD. Although preliminary research shows that attempts to modify or improve reasoning can affect change in paranoia (Garety et al., 2015), a more precise understanding of the environmental variables that directly affect PD and related phenomena may yield more precise and effective interventions.

In other words, if a researcher's goal is to predict behavioral *effects* under certain environmental conditions (as is often the case for those working at the mental level), then cognitive biases, ToM deficits, defensive efforts, or any other mental or non-mental variable (e.g., genetic or neurological activity) can be used to achieve prediction, so long as it reliably precedes that effect. If, however, a researcher's goal is to *predict-and-influence* behavior (a key goal in CBS) then this places certain constraints on the types of analyses they are ultimately interested in. If they want to influence behavior then they must identify variables that allow them to achieve this goal. Given that mental mechanisms can never be directly manipulated, such researchers turn to environmental variables that can be manipulated and that are functionally related to the behavior of interest. This not only improves our

understanding of factors that give rise to and maintain delusional beliefs but also occasions better influence over those beliefs.

Part II: CBS, RFT, and AARR

Contextual Behavioral Science (CBS)

CBS is an empirically and theoretically productive intellectual tradition rooted in a philosophy of science known as 'functional contextualism'. CBS draws upon a limited number of behavioral principles and analytic concepts in order to understand, predict, and influence behavior, in ways that are precise (a limited number of analytic concepts are used), high in scope (an analytic concept applies to a range of cases) and which have depth (analytic concepts cohere across levels of analysis) (Hayes, Barnes-Holmes, & Wilson, 2012). Within CBS, behavior is considered in its historical and situational context. The context affecting such behaviors ('acts-in-context') can be proximal (e.g., eye contact making me look away) and/or distal (e.g., traumatic experiences two years ago influence whether I socialize today).

For the reasons noted in Part I, CBS makes no appeal to mediating mental mechanisms. Instead it explains behavior in terms of environmental events and does so by identifying functional relations between (past and present) environment and behavior. To illustrate, consider JTC and PD. Although the correlation between JTC and PD is interesting, CBS researchers would not stop there. They would also investigate what this or other related patterns of behavior are a *function of* (i.e., what history of learning and environmental factors give rise to and maintain JTC). Once identified, these environmental variables can be manipulated to exert influence over the behavior of interest. For instance, by reinforcing efforts to seek out additional information before drawing conclusions we may improve performance on probabilistic reasoning paradigms and influence JTC in (real-life) situations. Hence, in CBS the study of delusional beliefs deemphasizes mental mediators (e.g., JTC,

attentional biases) and instead searches for environmental moderators (e.g., antecedents and consequences that give rise to and maintain delusional behaviors).

Relational Frame Theory

Relational Frame Theory (RFT; Hayes et al., 2001) is a key CBS account that posits that a type of operant behavior known as arbitrarily applicable relational responding (AARR) emerges relatively early on in human development and underpins human language and cognition. *Relational responding* refers to responding to one stimulus in terms of another and RFT distinguishes two types. Many species can respond based on the physical relationship between stimuli (e.g., picking something physically bigger or smaller than something else), which is referred to as non-arbitrarily applicable relational responding (NAARR) because the physical or non-arbitrary properties of the stimuli being related determine the relational response. However, humans also learn to relate stimuli regardless of their physical properties and in ways never experienced or instructed in the past (i.e., to *derive* relations between stimuli). RFT refers to this behavior as arbitrarily applicable relational responding (AARR) because it is based on contextual cues independent of the physical properties of the stimuli being related. For example, if I tell you that A is bigger than B, then you can derive that B is smaller than A. This relational response is not controlled by the physical properties of A and B but rather the words 'bigger' and 'smaller'.

AARR is defined in terms of three core properties: (i) *Mutual entailment* refers to the derived bi-directionality of stimulus relating such that if A is related to B in a specific context, then a relation from B to A is also entailed in that context (e.g., if A > B then B < A). (ii) *Combinatorial entailment* refers to the combination of two or more relations. For example, if A is related to B and B is related to C, then a bidirectional relation between A and C is entailed (e.g., if A > B > C, then A < C and C > A). (iii) *Transformation of stimulus functions* refers to the finding that when stimuli are related - and the stimulus functions

("psychological properties") of one of those stimuli is modified in some way - the corresponding functions of related stimuli can change without additional training or instruction (e.g., if A < B < C and A is threatening, then C is more threatening than A).

Humans can learn to relate stimuli in many ways, from equivalence and opposition, to comparison, temporality, conditionality, hierarchy, analogy and deixis (or perspective-taking) (see Hughes & Barnes-Holmes, 2016a). The way in which people relate stimuli and the functions transformed through those relations is controlled by stimuli ('contextual cues') in the environment. Some ('relational') cues (e.g., the words "same as" or "bigger than") specify how stimuli are related while other ('functional') cues specify the functions transformed through those relations. For example, in the phrase "Zim tastes like chocolate", the relational cue 'like' specifies a sameness relation between the novel word (Zim) and chocolate while the functional cue 'tastes' specifies that the functions that should transform from chocolate to Zim are gustatory (and not visual or auditory).

While mutual and combinatorial entailment and transformation of functions represent the core defining features of AARR, there are three dimensions of analysis that are worth noting here, particularly when considering the role of AARR in psychopathology. The first is that of *complexity*: the number and type of (a) stimuli and ways in which they are related, (b) functions that are transformed through those relations, and (c) sources of contextual control that influence the probability of those responses can vary from higher to lower complexity. The second is the extent to which a particular pattern of relational responding has been *derived* previously, which can vary along a continuum from 'low' to 'high'. For example, if an individual is informed that B is the same as A and that C is the same as B they will likely derive that A is the same as C. The first instance in which they derive the A-C relation could be considered a 'highly derived' response given that it is entirely derived from a limited set of prior learning experiences. However, given an ever increasing number of opportunities to

derive the relation between those same stimuli (i.e., A and C), across each of the following occasions of derivation the resulting response may come to be increasingly defined as a 'low derivation' response (Hughes, Barnes-Holmes, & Vahey, 2012). The same may hold true for increasingly complex relational responses (sometimes referred to in non-technical terms as 'beliefs'). For instance, a belief derived for the first time may differ from a belief derived 100 or 1000 times. A third dimension along which relational responses can be analyzed is *coherence*, "the extent to which a particular pattern of relational responding yields relatively consistent consequences" (Hussey, Barnes-Holmes, & Barnes-Holmes, 2015, p. 13). AARR is an operant behavior, and as such, can be shaped by contingencies of reinforcement or punishment. Early on in our development these contingencies are delivered by the socio-verbal community to ensure that people AARR in an "internally consistent" or coherent manner. Derived relational responses coherent with previous learning will likely be reinforced by others while incoherent responses may be punished. These interactions will likely occur across a wide variety of stimuli and contexts, delivered via a host of different individuals, so that eventually coherence is a conditioned reinforcer for AARR[2].

Why is AARR so important? CBS researchers are interested in AARR for several reasons. First, learning how to relate stimuli, unconstrained by their physical properties, unshackles humans from the physical world. Now (in principle) anything can be 'symbolically' related to anything else. Indeed, once a sophisticated repertoire of AARR is established, then the ability to arbitrarily relate stimuli in even just a few ways quickly allows people to respond in a complex manner (i.e., to generate relational networks under the control of contextual features specifying multiple relations and functions). Moreover, stimuli participating in derived relations can acquire entirely new functions or have their existing functions modified. This learned capacity to transform the functions of stimuli through

[2] For a more detailed treatment of complexity, derivation, and coherence see Hughes et al. (2012) or Barnes-Holmes, Barnes-Holmes, Hussey, & Luciano (2016)

derived relations appears to be the defining property of human language and cognition and responsible for its complexity, generativity, and semantic richness. Second, AARR unlocks a functional-analytic account of a host of complex behaviors such as thinking, self-knowledge, problem-solving, and rule following. For instance, Barnes-Holmes and colleagues (2001) suggested that higher level cognition requires the "ability to elaborate entire networks of stimulus relations quickly, to bring them under increasingly subtle forms of contextual control, to transform stimulus functions through entire networks, and to abstract entire features of the natural environment what will support and sustain relational responding" (p. 161). Thus, AARR equips humans with "an extraordinarily efficient and generative means of interacting with the world around them" (Hughes & Barnes-Holmes, 2016b, pp. 184-185). Third, AARR provides useful units with which to conceptualize 'private' behavior in functional-analytic terms, thus opening up such behavior to prediction-and-influence (e.g., via manipulating contextual control to produce changes in AARR). In this way, a restricted set of behavioral principles and analytic concepts may facilitate prediction-and-influence over an array of complex human behavior in a coherent and parsimonious manner. Over the past two decades, RFT research has expanded rapidly into many areas of important human concern. One such area is psychopathology[3].

The Consequences of AARR

AARR enables humans to respond to and modify their own behavior in ways that many other organisms cannot. Humans can link past with future events as well as derive entirely new sets of affairs from a limited set of prior experiences. Yet, AARR can also set the stage for many psychological problems including depression, anxiety, addiction, and suicidal ideation (see Törneke, Luciano, & Valdivia Salas, 2008). Regarding PD, an individual who had a traumatic experience (e.g., being betrayed) may find recalling this event

[3] For communicative purposes we have presented a simplified version of the origins, properties, and implications of AARR. For a far more nuanced perspective we encourage readers to consider Hayes et al. (2001); Hughes and Barnes-Holmes (2016a, b), Stewart and Roche (2013) or Törneke (2010).

highly aversive because the words used to describe those events participate in bidirectional relations with the actual events, and, as a result, acquire many of the latter's functions. AARR can also put people into contact with purely hypothetical events. For example, a hypothetical future event involving being betrayed can be contacted via temporal relations (e.g., *"If I trust someone again they will betray me"*). Although this event has not, and may never, happen it can nevertheless elicit fear and distress via derived stimulus relating. Indeed, AARR may explain how people come to fear and avoid certain stimuli in the absence of prior experience or instruction (see Part III).

From Simplicity to Complexity: Stimulus Relations to Rules

Within CBS, increasingly complex instances of AARR have typically been studied under the rubric of 'rule-governed behavior'. As previously mentioned, stimuli can be related in many ways using relational and functional cues. A rule is a complex stimulus relation involving "the coordination of behavior with a verbally specified contingency that makes reference to antecedent, behavioral, and consequential events" (Hughes & Barnes-Holmes, 2016b, p. 195). In other words, when humans (capable of AARR) generate and follow rules they are responding to the relations between the stimuli contained in those rules. Imagine, for example, that a person with PD says to themselves, "*If I go outside then people might try to attack me but if I stay inside I'll be safe*". This rule consists of a number of stimuli (I vs. other people) and events (staying inside vs. going outside, being attacked vs. being safe) that are related in a specific manner. The conditional ('If', 'Then') and temporal ('Then') relations specify the order of events and stipulate the behavior required (staying indoors) in order to avoid the aversive consequences of going outside (being attacked). The listener can determine if the rule is being followed and if it 'works' by assessing whether their behavior coordinates with the relations specified (i.e., are they staying inside?) and results in the specified consequences (i.e., safety).

Rule-following is a double-edged sword. It allows us to set and achieve goals, profit from other's experience and even deal with events before they occur (e.g., one can follow the rule "*Don't walk through that neighborhood at night, many people have been mugged there"* and avoid harm without having to directly contact the events involved). Yet, rule-following may also set the stage for human suffering. Research indicates that following rules (whether self- or socially-generated) can often undermine a person's sensitivity to environmental changes, which may lead to undesirable consequences. For example, when given instructions about how to complete a task, people often become insensitive to subsequent changes in experimental contingencies, even when these changes render the original rule ineffective (e.g., Doll, Jacobs, Stanfey, & Frank, 2009; Hayes, Brownstein, Zettle, Rosenfarb, & Korn, 1986; Wulfert, Greenway, Farkas, Hayes, & Dougher, 1991). Furthermore, this "insensitivity" effect may be amplified for sub-clinical and clinical conditions, including dysphoria (e.g., Baruch, Kanter, Busch, Richardson, & Barnes-Holmes, 2007), depression (e.g., McAuliffe, Hughes, & Barnes-Holmes, 2014), excessive gambling (e.g., Dixon, Hayes, & Aban, 2000), hallucination- and delusion-proneness (e.g., Cella, Dymond, & Cooper, 2009), and schizophrenia with delusions (e.g., Monestes, Villatte, Stewart, & Loas, 2014).

Part III: A CBS Approach to PD

We propose that (persecutory) delusional beliefs can be defined functionally as behaviors, and in particular, as AARR. Derived stimulus relating may explain why PD sufferers respond with fear, anxiety, and worry, or even attempt to escape or avoid particular stimuli and events, despite having not encountering them previously. Increasingly complex instances of AARR (rules) may also be central in the development and maintenance of PD, further restricting the individual's behavioral repertoire, while efforts to respond in ways that are coherent with those rules may increase their influence. A common thread through each of these topics is deictic (perspective-based) relating, while simultaneously, the extent to which

responses have been previously derived may help explain why many PD behaviors are emitted "automatically".

AARR: Fear and Avoidance

One could try to explain PD behaviors using a direct contingency learning approach. For instance, individuals with psychosis consistently report childhood traumatic events, including physical, sexual, and emotional abuse (e.g., Bondoli et al., 2013; Varese et al., 2012). Perhaps threat anticipation and related safety behaviors exhibited by some PD sufferers are the product of repeated stimulus pairings (e.g., adults paired with physical or emotional pain) or the relationship between past behavior and its consequences (e.g., avoidance of parents or others made aversive consequences less likely during childhood). These past experiences may subsequently influence the extent to which other (unknown) individuals are avoided in the present. Yet, this account faces several problems. First, it necessitates a history of traumatic experiences: but not everyone with PD has such a history, nor does everyone with such a history develop PD. Second, it fails to explain why certain individuals are avoided while others are readily approached. One could appeal to stimulus generalization, if the to-be-avoided person is similar to an original persecutor. Yet PD sufferers fear individuals dissimilar from the original persecutor. Indeed, delusional beliefs often incorporate imagined persecutors with no apparent relation to the previous persecutor (e.g., the FBI). Third, many cases of PD are deeply intertwined with language or symbolism. For example, the individual may interpret innocuous words as codes used by perceived persecutors to communicate to or about them, or may respond to the verbal and non-verbal behavior of others as evidence of threat (Anonymous, 2011; Adam, 2011). Hence, while direct contingency learning may play a role in PD, it cannot offer a sophisticated account by itself.

The fact that prior (traumatic) experience and stimulus generalization alone are insufficient to explain PD may have sparked interest in the aforementioned cognitive accounts. This reflects a growing consensus that verbal (symbolic) processes are crucial in the acquisition and maintenance of clinical behavior. This realization is consistent with RFT research suggesting that AARR can alter how people respond to stimuli in the absence of direct experience or instruction. More specifically, people can come to fear or avoid stimuli not based on direct experiences with them, but due to how those stimuli are related. For instance, Dymond, Roche, Forsyth, Whelan, and Rhoden (2008) taught people that two arbitrary stimuli (B1 and C1) were the same and that both of these stimuli were opposite to two other stimuli (B2 and C2). Participants subsequently learned that pressing one button (when shown B1) allowed them to avoid an unpleasant stimulus while pressing another button (when shown B2) allowed them to avoid pleasant stimuli. People subsequently not only avoided stimuli directly (B1) and indirectly (C1) related to unpleasant events but also approached those that signaled 'safety' or the absence of such events (C2) (also see Dougher, Hamilton, Fink, & Harrington, 2007; Dymond, Schlund, Roche, DeHouwer, & Freegard, 2012; Roche, Kanter, Brown, Dymond, & Fogarty, 2008).

Drawing on these findings, we argue that AARR may also be responsible for the formation and generalization of fear, threat-beliefs, avoidance, and other psychological properties (e.g., worry) characteristic of PD, especially for stimuli and events never previously contacted. Dymond, Schlund, Roche, and Whelan (2014) illustrate this. Two groups of participants (subclinical spider-phobic and non-phobic individuals) were taught two equivalence relations (A1-B1-C1 and A2-B2-C2). They then learned that one of these stimuli (B1) was "threatening" because it signaled the presentation of a spider while another (B2) was "safe" because it signaled removal of a spider. Both groups avoided stimuli directly (B1) and indirectly (C1) related to presentation of a spider; however the spider-phobics showed

higher avoidance than the non-phobics. It also took the phobic group fewer trials to learn to avoid the spider-related stimulus. This raises several possibilities. First, the observed differences between the spider-phobic and control groups may have been due to relative differences in their histories of *derivation* with respect to spider stimuli. Second, spiders may also have had different functions for the phobic group (e.g., fear) compared to controls (e.g., ambivalence, curiosity, disgust). Extrapolating to PD suggests that socially-relevant stimuli may have particular functions (e.g., danger, threat, etc.) for PD sufferers that they don't for others. Third, individuals with PD may have a longer history of AARR whereby they respond to social stimuli (e.g., faces) as potentially harmful, which might explain functional differences in responding to social stimuli between those with and without PD, as well as the speed with which social stimuli occasion fear and avoidance in this population.

Hence, AARR may explain how people come to fear, or try to escape or avoid, certain stimuli despite never encountering them previously. Fear, worry, threat anticipation, escape, and avoidance in PD may all be underpinned by similar processes. This possibility - that derived relating is central to PD - certainly warrants further investigation. Likewise, Dymond et al.'s (2014) findings raise the possibility that the number of times a particular relation has been derived might influence the ways people respond to novel but related stimuli. Regarding PD, it may be that once an aversive interpersonal event is experienced, derived stimulus relating allows that event to be placed into contact with entirely unrelated events, which transforms the functions of those events as well. Once someone can AARR, an aversive event (e.g., approaching a stranger) need not be directly experienced for someone to respond to it. They may formulate a network of related events (rules) characterized by fear and avoidance functions (e.g., "*approaching strangers is dangerous, therefore I should avoid people I don't know*"). Furthermore, as these relational networks are increasingly derived, the fear and avoidance responses they occasion may come to be increasingly elicited "automatically".

Rule-Governed Behavior

As outlined, rule-following can help us navigate the world effectively. But it also has a "dark-side", undermining sensitivity to environmental changes, which can lead to aversive consequences or missing out on appetitive ones. Rule-following may be central to the development and persistence of delusional beliefs. For instance, delusional beliefs can be considered as *rules* given that beliefs usually specify an antecedent (e.g., *stranger*), behavior (e.g., *trying to harm me*) and consequence (e.g., *if I leave the house*). To illustrate, consider the work of Monestes et al. (2014) who exposed two groups (patients formerly presenting with delusions and controls) to a simple learning task. Participants encountered a particular set of contingencies and then, unknown to them, the contingencies were reversed. They were either given a rule about the original contingencies, asked to generate a rule for themselves, or given no information. Unlike the control group, many of the patients "stuck" with their learning about the initial contingences and did not adjust. This was most prevalent in patients provided with a rule or asked to devise one themselves. Patients not given or asked to generate rules adjusted far more readily. Hence, rule-governed behavior can undermine contact with environmental contingencies. Regarding PD, this may suggest why delusional beliefs can be resistant to extinction (as earlier learning experiences come to dominate behavior, making adjustment to novel contingencies difficult). Further research into the relationship between PD and rule-based insensitivity is needed, perhaps replicating this paradigm with individuals with PD versus controls. This work could also be extended to examine moderators of rule-following in PD, from the nature of the antecedent, response, or consequence specified by the rule, to the rule's source (e.g., self vs. social), the believability of the rule or rule giver, and whether the rule was initially accurate or inaccurate.

Individuals with PD develop multiple rules for navigating their physical, social, and verbal worlds, including rules about themselves as vulnerable, others as malicious, and the

world as dangerous. Their behavior also appears governed by self-generated rules designed to escape or avoid aversive events (e.g., perceived interpersonal harm). Such avoidance is likely negatively reinforced via short-term consequences of responding in-line with the rule (e.g., avoiding others leads to absence of harm and reduces anxiety). However, as Freeman et al. (2002) note, engaging in such behaviors may prevent contact with other contingencies providing evidence contrary to the rule (e.g., appetitive interpersonal events). Thus, rule-governed behavior characterized by insensitivity to contingencies, and which falls under control of short-term consequences, may be important in the development and maintenance of PD. As new PD-relevant rules are generated, and new events experienced, this relational network may grow, perhaps dominating over other sources of behavioral regulation. This may further limit the behavioral repertoire and explain why PD behaviors are often resistant to extinction. Thus the idea of persecutory thoughts and ideas being elaborated to form a persecutory belief system can be understood in CBS terms as relational networks involving PD-related stimuli, rules, and transformation of functions through those networks.

Motivative Augmentals

Michael (1982) defined ‘motivating (or establishing) operations’ (MOs) as “any change in the environment which alters the effectiveness of some object or event as a reinforcer and simultaneously alters the momentary frequency of the behavior that has been followed by that reinforcement” (p. 151). Within PD fear may function as a motivating factor. Michael (1982) conceptualized fear (in operant terms) as an increased capacity for one's responses to be reinforced by removal of certain stimuli and increased frequency of behavior achieving their removal. This offers a useful starting point for understanding fear and threat responding in PD. However, in CBS, we extend the idea of MOs to events whose motivational functions have been acquired by means other than prior experience; namely, via transformation of functions (Ju & Hayes, 2008). In RFT, the term ‘augmenting’ has been

used to refer to "rule-governed behavior due to relational networks that alter the degree to which events function as consequences" (Barnes-Holmes et al., 2001, p. 109). Behavior due to relational networks that momentarily alter the degree to which previously established consequences function as reinforcers or punishers is called 'motivative augmenting'. A 'motivative augmental' is a relational network that can regulate behavior by transforming the motivating functions of a stimulus or event, even though a particular set of consequences might never have been contacted before (Valdivia, Luciano, & Molina, 2006). Within PD, an individual may generate a motivative augmental such as *"Other people are evil – they would hurt me if given a chance"*, even though they may have never experienced physical harm from others. This rule may augment the aversive functions of encountering known and unknown others. Words such as "evil" and "hurt" present sensory, emotional, and perceptual functions of potential interpersonal threat, and, through derived relations, other people may acquire these functions. Hence the individual may be more likely to avoid strangers (e.g., they might get off a bus early following an ambiguous event, such as being stared at).

Wray, Freund, and Dougher (2009) postulated that motivative augmentals may be implicated in the cognitive biases observed in clinical populations. They proposed an increased probability of behavior relevant to the contingencies potentiated by a particular motivative augmental with a decreased probability of behaviors relevant to other contingencies. For example, someone with PD acting under motivative augmental influence of this kind might attend to potentially threatening stimuli and evade social situations all the while avoiding other behaviors, such as interacting with others, thereby missing events contradicting that motivative augmental. In other words, motivative augmentals can increase the probability that a limited range of stimuli come to exert control over behavior, leading to a narrowing of the behavioral repertoire and a reduced likelihood of contacting other environmental contingencies. Motivative augmentals may also lead people to discriminate

certain stimuli quicker and more frequently than others. Ambiguous stimuli may be discriminated and related in ways consistent with the motivative augmental. This bears similarities to Freeman et al.'s (2013) view that, within PD, people will preferentially process threatening aspects of their environment, and that threatening interpretations of ambiguous events (e.g., looks from strangers) will be made. However, in RFT the functional relation between environment and behavior is specified thus potentially offering greater influence.

Relational Coherence

Coherence may also help to explain cognitive and attentional biases within PD and the persistence of PD despite contradictory evidence. Coherence has been conceptualized as "the behavior of generating a narrative that accurately, coherently, and contingently describes relations among events" (Wray, Dougher, Hamilton, & Guinther, 2012, p. 600). Responding coherently is socially reinforced early in development and is argued to serve as a conditioned reinforcer for relational responding (Roche, Barnes-Holmes, Barnes-Holmes, & Hayes, 2001). Responding coherently has two important consequences. First, discriminating that one is relating coherently (i.e., being "correct"), or that certain patterns of relational responding increase the likelihood of certain consequences (i.e., prevention of harm) may make similar (future) responses more likely. Second, the consequences that historically followed from responding coherently would suggest that, once established, relational networks may be quite difficult to break up or eliminate, even in the face of direct, contradictory training (Wilson & Hayes, 1996). These two consequences of responding coherently may help explain why delusional beliefs are particularly resistant to extinction.

Self- versus social-coherence. The difference between responding in ways that are coherent with one's own history of relating versus with how the wider social community relationally responds may be important in many clinical domains, especially PD. Munro (1999) proposed that the content of persecutory beliefs is logically constructed and internally

consistent, but incoherent with the responding of the community. Freeman et al. (2001; 2002) suggested that delusions can be viewed as an individual's attempt to make sense of anomalous events and that difficulty in tolerating ambiguity (and perhaps incoherence) may increase the likelihood that persecutory beliefs are accepted. Several CBS researchers have suggested that the reinforcing functions of coherence may explain why sense-making continues despite its aversive consequences (see Hayes, Strosahl, & Wilson, 1999; Wray et al., 2012). This may also explain why people often persist in ineffective rule-following despite changes in contingencies - and as such - why delusional beliefs are particularly resistant to extinction.

Consider a recent CBS study by Wray and colleagues (2012) who investigated whether coherence functions as a conditioned reinforcer for relational responding. They provided participants with solvable, unsolvable, and neutral tasks. In the solvable task the contingencies remained consistent and feedback was delivered contingent on correct and incorrect responding. In the unsolvable task feedback was yoked to participants' performance on the solvable task (presented first), independent of performance on the unsolvable task. Finally, the neutral tasks were formally similar to the solvable task but did not encourage sense-making, thus allowing investigation of whether participants prefer solvable conditions or simply to avoid unsolvable ones. Wray et al. then assessed preferences for each type of task. They found no strong preferences for solvable over neutral tasks, suggesting that people may have been motivated to avoid incoherent (unsolvable) tasks rather than engage in those that reinforced coherent responding (solvable tasks). Participant reports also suggested that coherence may serve different functions for different people and may influence some individuals more than others (see also Bordieri, Kellum, Wilson, & Whiteman, 2015). Regarding PD, perhaps individuals high on the continuum find incoherence more aversive than those lower on that same continuum.

Coherence in the Face of Ambiguity. Research suggests that individuals with PD experience a sense of danger and vigilance towards potential sources of threat in ambiguous contexts, as well as a belief, without evidence, that a threat exists (Phillips, Senior, & David, 2000). Recent RFT work on responding to ambiguous relational networks may be relevant. Quinones and Hayes (2014) exposed participants to a learning task wherein two different relational networks were trained. One (A1 < B1, A1 > C1) was designed to allow people to relate stimuli coherently while the other (A2 > B2, C2 < A2) did not allow for a single relationship to be derived between B2 and C2. Afterwards, participants were asked to indicate if the B and C stimuli were ‘similar’ or 'different’ from one another. In the first (unambiguous) network participants responded to B as different to C (i.e., coherently given that B1 > C1). However, in the second (ambiguous) network participants differed idiosyncratically in how they classified the B-C relation, some acting as if B and C were different and others as if they were similar. Interestingly, people treating B as different to C consistently *also* acted as if one stimulus was greater than the other in a later task while those treating B and C as similar did not.

Hence, relational coherence is evident even for ambiguous stimuli and events. These findings showcase how behavior may break down when relational networks do not cohere and how people strive to respond coherently in incoherent contexts. The authors also argued that these findings demonstrate how responding to ambiguity is influenced by prior patterns of relational responding and that this might provide an initial behavioral model of cognitive biases and errors such as JTC and confirmatory biases (also see Vitale, Barnes-Holmes, Barnes-Holmes, & Campbell, 2008). Thus, early work on coherence and responding to ambiguous relational networks suggests that coherence can act as a conditioned reinforcer for relational responding. Furthermore, when presented with incomplete information or ambiguous stimuli, people try to make sense of these events based on prior relational

responding. Wray et al.'s (2012) findings suggest that certain individuals may be even more motivated to avoid contexts where they must respond incoherently compared to contexts where they can respond coherently. Regarding PD: first, perhaps incoherence and ambiguity are more aversive for people with PD. Perhaps for them one error in judgment (e.g., concerning strangers' intentions) might lead to dire consequences. Second, the highly aversive functions of ambiguity (Freeman et al., 2013), together with relevant motivative augmentals, may also help explain why PD sufferers respond to ambiguity negatively or perceive it as evidence of threat.

Summary

In Part III we highlighted ways in which PD-related relational networks might develop and how seemingly innocuous stimuli (e.g., other people's words or gestures) may be incorporated into these networks and acquire aversive functions. Fear, threat, and avoidance may be established and transformed from one stimulus to another without prior experience, and despite physical dissimilarity among stimuli. Transformation of functions through complex relational networks may account for the internally coherent nature of delusional beliefs and how fear and avoidance may be readily emitted given certain (social) stimuli or ambiguous stimuli. Maladaptive rule-following may also be relevant for maintenance of delusional beliefs and may help clarify why these beliefs persist despite contradictory evidence. Rule-governed behavior can undermine sensitivity to environmental changes and several studies show clinical populations as particularly susceptible in this respect. Rule-following and motivative augmentals may explain how PD sufferers can develop narrow repertoires that perpetuate delusional beliefs and related responding.

Preliminary studies on the reinforcing properties of coherence with non-PD populations indicate that individuals are motivated to seek coherence and avoid incoherence, and will attempt to make sense given ambiguous relational networks. Given the threatening

nature of PD, and the verbally contacted consequences of failing to coordinate one's behavior with rules, incoherence and ambiguity may be more aversive for individuals with PD. Those with PD may develop a persecutory belief system (relational network), that is elaborate, has been derived frequently, and is internally coherent. The extent to which these relations have been derived before will likely impact the level of coherence, and render contradictory (incoherent) information aversive. Finally, when people engage in maladaptive safety behaviors (e.g., avoidance) that correspond with this relational network, and that are reinforced via coordination with self-generated rules, they will likely undermine contact with other contingencies, leading to narrow, inflexible repertoires, which perpetuate this vicious cycle. Thus, PD sufferers may be unable to re-evaluate belief veracity, the effectiveness of self-instructions, or long-term adaptiveness of their behavior.

Part IV: Future Directions

This paper set out to draw attention to an understudied phenomenon in CBS (delusional beliefs) as well as a burgeoning RFT literature that could inform a functional treatment. Although a functional analysis of PD remains to be conducted, concepts such as coherence, complexity, and derivation, transformation of motivative functions, and rule-governed behavior will likely be key. Although empirical work on these is only beginning and little is known about their role in PD, we believe they may guide future research on this topic.

One immediate recommendation is that researchers examine the transformation of fear, threat, avoidance and other relevant functions (e.g., worry) via contingency learning and derived stimulus relating in PD populations. This would extend current literature and shed light on AARR in PD specifically. Second, researchers could examine how specific properties of AARR, including complexity, derivation, and coherence, impact elicitation of fear and avoidance by PD-related stimuli. For example, perhaps individuals with PD have

longer histories of fear and avoidance responding to socially-relevant stimuli (e.g., faces) and may demonstrate relatively higher fear responding to novel but related stimuli (e.g., new faces). Future studies could establish a series of relations between stimuli with threatening or fearful functions and arbitrary stimuli in both PD and non-PD sufferers. During training several participants could receive relatively few opportunities to derive these relations whereas others could receive many opportunities to do so. Researchers could then examine whether the number of times derived relational responses were emitted impacts the speed and accuracy of those responses (as well as the transformation of function). Others could examine if an increased history of derivation influences the extent to which arbitrary stimuli are automatically treated as threatening or dangerous as well as probability of avoidance given those stimuli. Regarding rule-governed behavior, future research could examine how PD correlates with rule based insensitivity using procedures similar to Monestes et al. (2014) and modified versions that include valenced stimuli directly pertinent to PD-relevant content.

More research is needed on coherence and responding in ambiguous social situations as these seem particularly relevant in PD (e.g., Freeman et al., 2002; Freeman, 2007; Phillips et al., 2000). For example, do incoherence and ambiguity have more aversive functions for PD sufferers? This could be assessed as in Quinones and Hayes (2014), capturing physiological data (e.g., heart rate, galvanic skin response) as well as self-report measures of tolerance for ambiguity and incoherence. This may help us understand why this population experience greater anxiety and perceive threat in ambiguous social contexts.

Two final points are noteworthy. First, deictic ('perspective-taking') relating is likely important in everything discussed. Take, for instance, rule-following: perhaps *self-generated* rules concerning antecedents, behaviors, and consequences of delusions differ from those of society. Perhaps the speed with which the former are generated and followed, the strength of the functions transformed, and their probability of being extinguished differ compared to the

latter. Now consider coherence: individuals may strive to respond in ways that are coherent with their self-generated rules more than for socially delivered rules. As previously discussed, PD involve the belief that (a) one is being persecuted and (b) that another person is purposefully responsible for this harm, and as such interpersonal deictic relating seems critical. Intrapersonal deictic responding (taking perspective on one's own responding) is also important, as capacity for re-evaluation of one's beliefs may also be deficient in PD (e.g., Garety & Freeman, 1999); we will touch on this again in the 'Clinical Implications' section.

Second, there are now tools that enable researchers to capture relational responses that have been extensively derived and that vary in respective complexity. Tools such as the Implicit Relational Assessment Procedure (IRAP; Barnes Holmes et al., 2010), Relational Responding Task (RRT; DeHouwer et al., 2015) and others (see Gawronski & DeHouwer, 2014) may shed light on behaviors emitted under so-called conditions of automaticity (i.e., quickly and without awareness, intention, or control). Predicting and influencing such behaviors may help refine prediction and influence of PD.

Clinical Implications

AARR is both a boon and a bane. Regarding PD, it facilitates core properties of PD; however it can also be used to undermine PD. For instance, via AARR, people can discriminate and modify their own behavior via verbal reports of its effectiveness. Some therapeutic approaches (e.g., Acceptance and Commitment Therapy [ACT]; Hayes et al., 1999) attempt to disrupt or modify functions of clinically-relevant experiences by changing the contexts in which they are usually related to behavior. They also encourage contact with other environmental contingencies so that individuals may alter and adapt their behavior accordingly. ACT clarifies that it is not PD responses themselves that are problematic but how people respond to them (e.g., whether they can take perspective on them). Such approaches may help PD sufferers interrupt the cycle of responding to PD-related events,

increase contact with direct contingencies ('disconfirmatory evidence'), and broaden their behavioral repertoires (see e.g., Morris, Johns, & Oliver, 2013).

While attempts to improve contact with other environmental contingencies might also produce modifications to existing PD-relevant networks (e.g., from *"the world is evil and people will try to exploit me"* to *"some people might try to exploit me but I can identify when this might happen through my experiences with others"*), unlike interventions that focus on challenging or altering 'cognitions' or 'schemas', CBS researchers make no appeal to mental mechanisms to alter behavior. However, such attempts to improve contact with other contingencies may be complicated. First, PD sufferers may have an extensive history of maladaptive rule-following and a limited behavioral repertoire and so additional (e.g., social) skills training may be required. Second, because not following PD-relevant rules may be perceived as disastrous, the idea of not doing so may be highly anxiety-provoking. Indeed, Hayes et al. (1999) argue that excessive or problematic rule-following aimed at reducing contact with aversive private events is a core feature of clinical phenomena. This is also in keeping with Freeman et al.'s (2002) view that anxiety is associated with PD and their maintenance. From a CBS perspective, however, anxiety and other affective processes are not considered to have an ultimate causal role in maintaining avoidance behaviors. Behavior-behavior relations (e.g., feeling anxious and engaging in avoidance) may indeed be involved in clinically-relevant behavioral repertoires; however, a behavior-behavior relation is not a complete explanation. They are themselves based on contextual factors (e.g., contingencies of reinforcement; Hayes & Brownstein, 1986), that can be manipulated to achieve influence over behavior[4].

[4] It is important to note that many of the terms we have referred to in this paper, such as coherence, rule-governed behavior, motivative augmentals, etc., are non-technical terms (Barnes-Holmes, Hussey, McEnteggart, Barnes-Holmes & Foody, 2016). Furthermore, an RFT analysis of terms such as complexity, derivation, coherence, and ambiguity is limited at present (see Barnes-Holmes et al., 2016). Thus, further research is needed to test and refine the assumptions outlined here. As such, a technical treatment of these phenomena will be needed before we can truly extrapolate them to the study of PD.

General Conclusion

Cognitive researchers have argued that the emergence and persistence of PD is linked to reasoning biases, affective responses (e.g., anxiety), and safety behaviors. Although this literature has highlighted important behavioral phenomena and environmental events pertinent to PD, it has mainly focused on the mental level of analysis and developing mental models in the service of prediction. Drawing on RFT research, we provide a bottom-up, functional-analytic interpretation of PD and identify a number of behavioral processes possibly involved. By defining PD in terms of AARR and by attempting to identify the learning histories and environmental factors that support PD-related behavior, it may be possible to not only predict PD phenomena, but to influence them. However, before such a functional analysis can take place additional research is needed on key concepts (e.g., AARR properties such as complexity, derivation). Non-technical terms and concepts (e.g., rule-governed behavior, coherence) must be studied and refined. Although our account is based on behavioral principles and existing research, ultimately data will drive the functional analysis needed. We hope our interpretation may provide some initial guidance on this path.

References

American Psychiatric Association (2000). *Diagnostic and Statistical Manual of Mental Disorders* (4th Ed. text rev.). Washington: Author.

American Psychiatric Association (2013). *Diagnostic and Statistical Manual of Mental Disorders* (5th Ed.). Arlington, VA: American Psychiatric Publishing.

Adam (2011). Experiencing suspicious thoughts and paranoia: An account. *Schizophrenia Bulletin, 37,* 656–658.

Anonymous (2011). Language games, paranoia, and psychosis. *Schizophrenia Bulletin, 37,* 1099–1100.

Barnes-Holmes, D., Barnes-Holmes, Y., Hussey, I., & Luciano, C. (2016). Relational Frame Theory. Finding its historical and intellectual roots and reflecting upon its further development: An introduction to Part II. In S. Hayes, D. Barnes-Holmes, R. Zettle, and T. Biglan (Eds.), *Handbook of Contextual Behavioral Science* (pp. 117-128). West Sussex, UK: Wiley-Blackwell.

Barnes-Holmes, Y., Hussey, I., McEnteggart, C., Barnes-Holmes, D., & Foody, M. (2016). Scientific ambition: The relationship between Relational Frame Theory and middle-level terms in Acceptance and Commitment Therapy. In R. D. Zettle, S. C. Hayes, D. Barnes-Holmes & A. Biglan, (Eds.) *The Wiley Handbook of Contextual Behavioral Science* (pp. 365-382). West Sussex, UK: Wiley-Blackwell.

Barnes-Holmes, D., O'Hora, D., Roche, B., Hayes, S. C., Bissett, R. T., & Lyddy, F. (2001). Understanding and verbal regulation. In S. C. Hayes, D. Barnes-Holmes, & B. Roche, (Eds.) *Relational Frame Theory: A Post-Skinnerian Account of Human Language and Cognition* (pp. 103-118). New York: Plenum.

Baruch, D. E., Kanter, J. W., Busch, A. M., Richardson, J. V., & Barnes-Holmes, D. (2007). The differential effect of instructions on dysphoric and nondysphoric persons. *Psychological Record, 57,* 543-554.

Beards, S., & Fisher, H. L. (2014). The journey to psychosis: An exploration of specific psychological pathways. *Social Psychiatry & Psychiatric Epidemiology, 49,* 1541-1544.

Bebbington, P., McBride, O., Steel, C., Kuipers, E., Radovanovic, M., Brugha, T., Jenkins, R., Meltzer, H., & Freeman, D. (2013). The structure of paranoia in the general population. *British Journal of Psychiatry, 202,* 419–427.

Bentall, R. P. (2006). Madness explained: Why we must reject the Kraepelinian paradigm and replace it with a 'complaint-orientated' approach to understanding mental illness. *Medical Hypotheses, 66,* 220–233.

Bentall, R. P., Corcoran, R., Howard, R., Blackwood, N., & Kinderman, P. (2001). Persecutory delusions: A review and theoretical interpretation. *Clinical Psychology Review, 21,* 1143-1192.

Bentall, R. P., de Sousa, P., Varese, F., Wickham, S., Sitko, K., Haarmans, M., & Read, J. (2014). From adversity to psychosis: Pathways and mechanisms from specific adversities to specific symptoms. *Social psychiatry and psychiatric epidemiology*, *49*, 1011-1022.

Bentall, R. P., Kinderman, P., & Kaney, S. (1994). The self, attributional processes and abnormal beliefs: Towards a model of persecutory delusions. *Behaviour Research and Therapy, 32,* 331-341.

Birchwood, M., Iqbal, Z., Chadwick, P., & Trower, P. (2000). Cognitive approach to depression and suicidal thinking in psychosis: Ontogeny of post-psychotic depression. *British Journal of Psychiatry, 177,* 516-528.

Bonoldi, I., Simeone, E., Rocchetti, M., Codjoe, L., Rossi, G., Gambi, F., … Fusar-Poli, P. (2013). Prevalence of self-reported childhood abuse in psychosis: A meta-analysis of retrospective studies. *Psychiatry Research, 210*, 8–15.

Bordieri, M., Kellum, K. K., Wilson, K. G., & Whiteman, K. C. (2015). Basic properties of coherence: Testing a core assumption of Relational Frame Theory. *The Psychological Record, 66*, 83-98.

Cella, M., Dymond, S., & Cooper, A. (2009). Impairment in flexible emotion-based learning in hallucination- and delusion-prone individuals. *Psychiatry Research, 170*, 70-74.

Collip, D., Myin-Germeys, I., & Van Os, J. (2008). Does the concept of "sensitization" provide a plausible mechanism for the putative link between the environment and schizophrenia? *Schizophrenia Bulletin, 34*, 220-225.

Corcoran, R., Mercer, G., & Frith, C. D. (1995). Schizophrenia, symptomatology and social inference: Investigating "theory of mind" in people with schizophrenia. *Schizophrenia Research, 17*, 5-13.

De Houwer, J., Heider, N., Spruyt, A., Roets, A., & Hughes, S. (2015). The Relational Responding Task: Toward a new implicit measure of beliefs. *Frontiers in Psychology, 6*, 319.

Dixon, M. R., Hayes, L. R., & Aban, I. B. (2000). Examining the roles of rule-following, reinforcement, and pre-experimental histories of risk-taking behavior. *Psychological Record, 50,* 687-704.

Doll, B. B., Jacobs, W. J., Sanfey, A. G., & Frank, M. J. (2009). Instructional control of reinforcement learning: A behavioral and neurocomputational investigation. *Brain Research, 1299*, 74-94.

Dougher, M. J., Hamilton, D. A., Fink, B. A., & Harrington, J. (2007). Transformation of the discriminative and eliciting functions of generalized relational stimuli. *Journal of the Experimental Analysis of Behavior, 88,* 179-197.

Drury, V. M., Robinson, E. J., & Birchwood, M. (1998). 'Theory of mind' skills during acute episode of psychosis and following recovery. *Psychological Medicine, 28,* 1101-1112.

Dymond, S., Roche, B., Forsyth, J. P., Whelan R., & Rhoden, J. (2008). Derived avoidance learning: Transformation of avoidance response functions in accordance with relational frames of same and opposite. *The Psychological Record, 58,* 271-288.

Dymond, S., Schlund, M. W., Roche, B., DeHouwer, J., & Freegard G. P. (2012). Safe from harm: Learned, instructed, and symbolic generalization pathways of human threat-avoidance. *PLoS One, 7,* DOI:10.1080/17470218.2013.800124

Dymond, S., Schlund, M. W., Roche, B., & Whelan, R. (2014). The spread of fear: Symbolic generalization mediates graded-threat avoidance in specific phobia. *The Quarterly Journal of Experimental Psychology, 67,* 247-259.

Fine, C., Gardner, M., Craigie, J., & Gold, I. (2007). Hopping, skipping or jumping to conclusions? Clarifying the role of the JTC bias in delusions. *Cognitive Neuropsychiatry, 12,* 46–77.

Freeman, D. (2007). Suspicious minds: The psychology of persecutory delusions. *Clinical Psychology Review, 27,* 425–457.

Freeman, D., Bentall, R. P., & Garety, P. A. (2008). *Persecutory Delusions: Assessment, Theory and Treatment*. Oxford University Press.

Freeman, D., Dunn, G., Fowler, D., Bebbington, P., Kuipers, E., Emsley, R., Jolley, S., & Garety, P. A. (2013). Current paranoid thinking in patients with delusions: The presence of cognitive-affective biases. *Schizophrenia Bulletin, 39,* 1281–1287.

Freeman, D., & Garety, P. A. (2000). Comments on the content of persecutory delusions: Does the definition need clarification? *British Journal of Clinical Psychology, 39,* 407-414.

Freeman, D., & Garety, P. A. (2014). Advances in understanding and treating persecutory delusions: A review. *Social Psychiatry and Psychiatric Epidemiology*, *49*, 1179-1189.

Freeman, D., Garety, P. A., Bebbington, P. E., Smith, B., Rollinson, R., Fowler, D., Kuipers, E., Ray, K., & Dunn, G. (2005). Psychological investigation of the structure of paranoia in a non-clinical population. *British Journal of Psychiatry, 186*, 427−435.

Freeman, D., Garety, P. A., Fowler, D., Kuipers, E., Bebbington, P. E., & Dunn, G. (2004). Why do people with delusions fail to choose more realistic explanations for their experiences? An empirical investigation. *Journal of Consulting and Clinical Psychology, 72,* 671–680.

Freeman, D, Garety, P. A., & Kuipers, E. (2001). Persecutory delusions: Developing the understanding of belief maintenance and emotional distress. *Psychological Medicine*, *31*, 1293–1306.

Freeman, D., Garety, P. A., Kuipers, E., Fowler, D., & Bebbington, P. E. (2002). A cognitive model of persecutory delusions. *British Journal of Clinical Psychology*, *41*, 331–347.

Freeman, D., Garety, P. A., Kuipers, E., Fowler, D. Bebbington, P. E., & Dunn, G. (2007). Acting on persecutory delusions: The importance of safety seeking. *Behaviour Research and Therapy, 45,* 89-99. DOI: 10.1016/j.brat.2006.01.014

Freeman, D., Lister, R., & Evans, N. (2014). The use of intuitive and analytic reasoning styles by patients with persecutory delusions. *Journal of Behavioral Therapy & Experimental Psychiatry, 45*, 454–458.

Freeman, D., McManus, S., Brugha, T., Meltzer, H., Jenkins, R., & Bebbington, P. (2011). Concomitants of paranoia in the general population. *Psychological Medicine, 41*, 923–936.

Freeman, D., Startup, H., Dunn, G., Wingham, G., Cernis, E., Evans, N., Lister, R., Pugh, K., Cordwell, J., & Kingdon, D. (2014). Persecutory delusions and psychological well-being. *Social Psychiatry & Psychiatric Epidemiology, 49,* 1045–1050.

Frith, C.D. (1992). *The Cognitive Neuropsychology of Schizophrenia*. Psychology Press.

Frith, C. D., & Corcoran, R. (1996). Exploring 'theory of mind' in people with schizophrenia. *Psychological Medicine, 26*, 521-530.

Garety, P., Fowler, D., Kuipers, E., Freeman, D., Dunn, G., Bebbington, P., & Jones, S. (1997). London-East Anglia randomised controlled trial of cognitive-behavioural therapy for psychosis II: Predictors of outcome. *British Journal of Psychiatry, 171*, 420–426.

Garety, P. A., & Freeman, D. (1999). Cognitive approaches to delusions: A critical review of theories and evidence. *British Journal of Clinical Psychology, 38,* 113-154.

Garety P. A., & Freeman, D. (2013). The past and future of delusion research: From the inexplicable to the treatable. *British Journal of Psychiatry*, *203*, 327–333.

Garety, P. A., Freeman, D., Jolley, S., Dunn, G., Bebbington, P. E., Fowler, D. G., & Dudley, R. (2005). Reasoning, emotions, and delusional conviction in psychosis. *Journal of Abnormal Psychology*, *114*, 373–384.

Garety, P. A., Freeman, D., Jolley, S., Ross, K., Waller, H., & Dunn, G. (2011). Jumping to conclusions: The psychology of delusional reasoning. *Advances in Psychiatric Treatment, 17*, 332-339.

Garety, P. A., Hemsley, D. R., & Wessely, S. (1991). Reasoning in deluded schizophrenic and paranoid patients: Biases in performance on a probabilistic inference task. *Journal of Nervous and Mental Disease, 179,* 194-201.

Garety, P. A., Waller, H., Emsley, R., Jolley, S., Kuipers, E. … Freeman, D. (2015). Cognitive mechanisms of change in delusions: An experimental investigation targeting reasoning to effect change in paranoia. *Schizophrenia Bulletin, 41*, 400-410.

Gawronski, B., & DeHouwer, J. (2014). Implicit measures in social and personality psychology. In H. T. Reis, & C. M. Judd (Eds.), *Handbook of research methods in social and personality psychology* (2nd edition) (pp. 283-310). New York: Cambridge University Press.

Greig, T. C., Bryson, G. J., & Bell, M. D. (2004). Theory of mind performance in schizophrenia: Diagnostic, symptom and neuropsychological correlates. *Journal of Nervous and Mental Disease, 192*, 12–18.

Hayes, S. C., Barnes-Holmes, D., & Roche, B. (2001). *Relational Frame Theory: A Post-Skinnerian Account of Human Language and Cognition*. New York: Kluwer Academic/Plenum Publishers.

Hayes, S. C., Barnes-Holmes, D., & Wilson, K. G. (2012). Contextual Behavioral Science: Creating a science more adequate to the challenge of the human condition. *Journal of Contextual Behavioral Science, 1*, 1-16.

Hayes, S. C., & Brownstein, A. J. (1986). Mentalism, behavior-behavior relations, and a behavior-analytic view of the purposes of science. *The Behavior Analyst, 9,* 175-190.

Hayes, S. C., Brownstein, A. J., Zettle, R. D., Rosenfarb, I., & Korn, Z. (1986). Rule-governed behavior and sensitivity to changing consequences of responding. *Journal of the Experimental Analysis of Behavior, 45*, 237-256.

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

Hughes, S., & Barnes-Holmes, D. (2016a). Relational Frame Theory: The Basic Account. In S. Hayes, D. Barnes-Holmes, R. Zettle, and T. Biglan (Eds.), *Handbook of Contextual Behavioral Science* (pp. 129-178). West Sussex, UK: Wiley-Blackwell.

Hughes, S., & Barnes-Holmes, D. (2016b). Relational Frame Theory: Implications for the Study of Human Language and Cognition. In S. Hayes, D. Barnes-Holmes, R. Zettle, and T. Biglan (Eds.), *Handbook of Contextual Behavioral Science* (pp. 179-226). West Sussex, UK: Wiley-Blackwell.

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

Hughes, S., DeHouwer, J., & Barnes-Holmes, D. (2016). The moderating impact of distal regularities on the effect of stimulus pairings: A novel perspective on evaluative conditioning. *Experimental Psychology, 63,* 20-44.

Hussey, I., Barnes-Holmes, D., & Barnes-Holmes, Y. (2015). From Relational Frame Theory to implicit attitudes and back again: Clarifying the link between RFT and IRAP research. *Current Opinion in Psychology, 2,* 11–15.

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

Kaney, S., & Bentall, R. P. (1992). Persecutory delusions and the self-serving bias. Evidence from a contingency judgment task. *Journal of Nervous and Mental Disease, 180*, 773–780.

Kaney, S., & Bentall, R. P. (1989). Persecutory delusions and attributional style. *British Journal of Medical Psychology, 62*, 191–198.

Kesting M. L., & Lincoln, T. M. (2013). The relevance of self-esteem and self-schemata in individuals with persecutory delusions. *Comprehensive Psychiatry, 54*, 766-789.

Korver-Nieberg, N., Fett, A-K., Meijer, C., Koeter, M., Shergill, S., Haan, L., & Krabbendam, L. (2013) Theory of mind, insecure attachment and paranoia in adolescents with early psychosis and healthy controls. *Australia New Zealand Journal of Psychiatry, 47*, 737-745.

Lyon, H. M., Kaney, S., & Bentall, R. P. (1994). The defensive function of persecutory delusions. Evidence from attribution tasks. *The British Journal of Psychiatry, 164*, 637–646.

McAuliffe, D., Hughes, S., & Barnes-Holmes, D. (2014). The dark side of rule-governed behaviour: An experimental analysis of problematic rule-following in an adolescent population with depressive symptomatology. *Behavior Modification, 38*, 587-613.

Mehl, S., Rief, W., Lüllmann, E., Ziegler, M., Kesting, M-L., & Lincoln, T. M. (2010). Are Theory of Mind deficits in understanding intentions of others associated with persecutory delusions? *Journal of Nervous and Mental Disease, 198*, 516- 519.

Monestes, J. L., Villatte, M., Stewart, I., & Loas, G. (2014). Rule-based insensitivity and delusion maintenance in schizophrenia. *Psychological Record, 64*, 329-338.

Michael, J. (1982). Discriminating between discriminative and motivational functions of stimuli. *Journal of the Experimental Analysis of Behavior, 37*, 149-155.

Moritz, S., & Woodward, T. S. (2006). The contribution of metamemory deficits to schizophrenia. *Journal of Abnormal Psychology, 115,* 15–25.

Moritz, S., Woodward, T. S., & Rodriguez-Raecke, R. (2006). Patients with schizophrenia do not produce more false memories than controls but are more confident in them. *Psychological Medicine, 36,* 659–667.

Morris, E. M. J., Johns, L. C., & Oliver, J. E. (2013). *Acceptance and Commitment Therapy and Mindfulness for Psychosis.* Wiley-Blackwell: John Wiley & Sons, Ltd.

Moutoussis, M., Williams, J., Dayan, P., & Bentall, R. P. (2007). Persecutory delusions and the conditioned avoidance paradigm: Towards an integration of the psychology and biology of paranoia. *Cognitive Neuropsychiatry, 12,* 495-510.

Munro, A. (1999). *Delusional Disorder: Paranoia and Related Illnesses.* Cambridge University Press.

Oltmanns, T. F. (1988). Approaches to the definition and study of delusions. In T. F. Oltmanns & B. A. Maher (Eds.), *Delusional Beliefs* (pp. 3-12). New York: Wiley.

Peters, E. R., Joseph, S. A., Day, S., & Garety, P. A. (2004). Measuring delusional ideation: The 21-item Peters et al. Delusions Inventory (PDI). *Schizophrenia Bulletin, 30,* 1005-1022.

Peters, E. R., Joseph, S. A., & Garety, P. A. (1999). Measurement of delusional ideation in the normal population: Introducing the PDI (Peters et al. Delusions Inventory). *Schizophrenia Bulletin, 25,* 553-576.

Phillips, M. L., Senior, C., & David, A. S. (2000). Perception of threat in schizophrenics with persecutory delusions: An investigation using visual scan path. *Psychological Medicine, 30,* 157-167.

Quinones, J. L., & Hayes, S. C. (2014). Relational coherence in ambiguous and unambiguous relational networks. *Journal of the Experimental Analysis of Behavior, 101,* 76-93.

Read, J., Fosse, R., Moskowitz, A., & Perry, B. (2014). The traumagenic neurodevelopmental model of schizophrenia. *Neuropsychiatry, 4,* 65-79.

Roche, B, Barnes-Holmes, D., Barnes-Holmes, Y., & Hayes, S. C. (2001). Social processes. In, S. C. Hayes, D. Barnes-Holmes, & B. Roche (Eds.). *Relational Frame Theory: A Post-Skinnerian Account of Human Language and Cognition* (pp. 197-210). New York: Kluwer Academic/Plenum Publishers.

Roche, B., Kanter, J. W., Brown, K. R., Dymond, S., & Fogarty, C. C. (2008). A comparison of "direct" versus "derived" extinction of avoidance. *The Psychological Record, 58,* 443-464.

Salvatore, G., Lysaker, P. H., Popolo, R., Procacci, M., Carcione, A., & Dimaggio, G. (2012). Vulnerable self, poor understanding of others' minds, threat anticipation and cognitive biases as triggers for delusional experience in schizophrenia: A theoretical model. *Clinical Psychology and Psychotherapy, 19*, 247-259.

Sartorius, N., Jablensky, A., Korten, A., Emberg, g., Anker, M., Cooper, J. E., & Day, R. (1986). Early manifestations and first-contact incidence of schizophrenia in different cultures. *Psychological Medicine, 16,* 909- 928.

Scherzer, P., Leveille, E., Achim, A., Boisseau, E., & Stip, E. (2012). A study of theory of mind in paranoid schizophrenia: A theory or many theories? *Frontiers in Psychology, 3*, 432.

So, S. H., Freeman, D., Dunn, G., Kapur, S., Kuipers, E., Bebbington, P., Fowler, D., Garety, P. A. (2012). Jumping to conclusions, a lack of belief flexibility and delusional conviction in psychosis: A longitudinal investigation of the structure, frequency, and relatedness of reasoning biases. *Journal of Abnormal Psychology, 121,*129-139.

Stewart, I. & Roche, B. (2013). Relational Frame Theory: An overview. In S. Dymond & B. Roche (2013). *Advances in Relational Frame Theory: Research and Applications* (pp. 51-72). Context Press.

Törneke, N. (2010). *Learning RFT: An Introduction to Relational Frame Theory and Its Clinical Application*. New Harbinger Publications, Inc.

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

Valdivia, S., Luciano, C., & Molina, F. J. (2006). Verbal regulation of motivational states. *The Psychological Record, 56,* 577-595.

Valiente, C., Cantero, D., Vázquez, C., Sanchez, Á., Provencio, M., & Espinosa, R. (2011). Implicit and explicit self-esteem discrepancies in paranoia and depression. *Journal of Abnormal Psychology, 120,* 691-699.

Varese, F., Smeets, F., Drukker, M., Lieverse, R., Lataster, T., Viechtbauer, W., … Bentall, R. P. (2012). Childhood adversities increase the risk of psychosis: A meta-analysis of patient-control, prospective- and cross-sectional cohort studies. *Schizophrenia Bulletin, 38,* 661–671.

Vitale, A., Barnes-Holmes, Y., Barnes-Holmes, D., & Campbell, C. (2008). Facilitating responding in accordance with the relational frame of comparison: Systematic empirical analyses. *The Psychological Record, 58,* 365-390.

Wilson, K. G., & Hayes, S. C. (1996). Resurgence of derived stimulus relations. *Journal of the Experimental Analysis of Behavior, 66,* 267-281.

Wray, A. M., Dougher, M. J., Hamilton, D. A., & Guinther, P. (2012). Examining the reinforcing properties of making sense: A preliminary study. *Psychological Record, 62,* 599-622.

Wray, A. M., Freund, R. A., & Dougher M. J. (2009). A behavior-analytic account of cognitive bias in clinical populations. *The Behavior Analyst, 32,* 29-49.

Wulfert, E., Greenway, D. E., Farkas, P., Hayes, S. C., & Dougher, M. J. (1994). Correlation between self-reported rigidity and rule-governed insensitivity to operant contingencies. *Journal of Applied Behavioral Analysis, 27,* 659-671.