



Swansea University  
Prifysgol Abertawe



## Cronfa - Swansea University Open Access Repository

---

This is an author produced version of a paper published in:

*The Behavior Analyst*

Cronfa URL for this paper:

<http://cronfa.swan.ac.uk/Record/cronfa34883>

---

### **Paper:**

Dymond, S., Bennett, M., Boyle, S., Roche, B. & Schlund, M. Related to Anxiety: Arbitrarily Applicable Relational Responding and Experimental Psychopathology Research on Fear and Avoidance. *The Behavior Analyst*

<http://dx.doi.org/10.1007/s40614-017-0133-6>

---

This item is brought to you by Swansea University. Any person downloading material is agreeing to abide by the terms of the repository licence. Copies of full text items may be used or reproduced in any format or medium, without prior permission for personal research or study, educational or non-commercial purposes only. The copyright for any work remains with the original author unless otherwise specified. The full-text must not be sold in any format or medium without the formal permission of the copyright holder.

Permission for multiple reproductions should be obtained from the original author.

Authors are personally responsible for adhering to copyright and publisher restrictions when uploading content to the repository.

<http://www.swansea.ac.uk/library/researchsupport/ris-support/>

Running head: FEAR AND AVOIDANCE

Related to Anxiety: Arbitrarily Applicable Relational Responding and Experimental  
Psychopathology Research on Fear and Avoidance

Simon Dymond<sup>1,2</sup>, Marc Bennett<sup>3</sup>, Sean Boyle<sup>4</sup>, Bryan Roche<sup>4</sup>, and Michael Schlund<sup>5</sup>

<sup>1</sup> Experimental Psychopathology Lab, Department of Psychology, Swansea University, Park  
Campus, Swansea, SA2 8PP, United Kingdom. Swansea University.

<sup>2</sup> Department of Psychology, Reykjavík University, Menntavegur 1, Nauthólsvík, 101  
Reykjavík, Iceland.

<sup>3</sup> Trinity College Institute of Neuroscience, Lloyd Building, Trinity College Dublin, Dublin 2,  
Ireland.

<sup>4</sup> Department of Psychology, Maynooth University, Maynooth, Co. Kildare, Ireland.

<sup>5</sup> Department of Psychology, University of Pittsburgh, Loeffler Building, Room 316, 121  
Meyran Avenue, Pittsburgh, PA 15213, USA.

[Accepted/In Press, 2017, *The Behavior Analyst*]

Address for correspondence: Simon Dymond PhD., Experimental Psychopathology Lab,  
Department of Psychology, Swansea University, Park Campus, Swansea SA2 8PP, United  
Kingdom.

**Conflict of interest:** The authors declare that they have no conflict of interest.

**Ethical approval:** This manuscript does not contain the findings of original research.

### Abstract

Humans have an unparalleled ability to engage in arbitrarily applicable relational responding (AARR). One of the consequences of this ability to spontaneously combine and relate events from the past, present and future may, in fact, be a propensity to suffer. For instance, maladaptive fear and avoidance of remote or derived threat may actually perpetuate anxiety. In this narrative review, we consider contemporary AARR research on fear and avoidance as it relates to anxiety. We first describe laboratory based research on the emergent spread of fear- and avoidance-eliciting functions in humans. Next, we consider the validity of AARR research on fear and avoidance and address the therapeutic implications of the work. Finally, we outline challenges and opportunities for a greater synthesis between behavior analysis research on AARR and experimental psychopathology.

*Keywords:* arbitrarily applicable relational responding; stimulus relations; generalization; experimental psychopathology; fear; avoidance; anxiety.

Over recent decades, research on arbitrarily applicable relational responding (AARR) has continued at pace, driven in large part by the potential to facilitate understanding and treatment of human suffering. Humans have a remarkable ability to engage in, and rely on, AARR. That is, when they are interacting relationally with their environment, humans are capable of such feats as landing vehicles on distant planets or shining lights into the abyss of dying stars. In fact, it may be impossible to find occasions when humans are not interacting relationally with the environment. And yet, despite the abundance of AARR and the advantages it confers, many of our species everyday fear invisible germs, social rejection or the catastrophic consequences of things that may never happen. Some of us, despite being in no physical pain, may be unable to get out of bed. A tendency to fear everything and to engage in excessive, maladaptive avoidance of both real and perceived danger is one of the obvious counterproductive consequences of this human propensity for AARR. Unfortunately, for many, this relational basis of human behavior leads logically and tragically to suffering and self-destruction. Yet, what makes these scenarios both possible and problematic is that they are capable of empirical scrutiny when approached in terms of AARR.

Arbitrarily applicable relational responding is a form of overarching or generalized operant behavior in which novel, untrained responses emerge from a subset of directly trained responses (Dymond & Roche, 2013). Research on AARR has shown that when language-able humans are taught a series of arbitrary interrelated conditional discriminations involving physically dissimilar stimuli, the stimuli involved often become related to each other in ways not explicitly trained (Sidman, 1994). The trained discriminations are said to be arbitrary because the stimuli used in training are not physically similar or related other than by the experimental procedures. To illustrate, if choosing Stimulus X in the presence of Stimulus A is taught (i.e., A-X), and choosing Stimulus Y in the presence of Stimulus A (i.e., A-Y) is also taught, it is likely that untrained relations will emerge between X and A, Y and A

(“symmetry”), X and Y, and Y and X (“equivalence”), in the absence of any feedback. When this occurs, an equivalence relation is said to have formed among the relata (Dymond & Roche, 2013; Hayes & Hayes, 1992). One of the reasons forms of AARR like equivalence relations are interesting is that they are not predicted by traditional behavioral principles of discrimination and generalization (Dymond & Rehfeldt, 2000); thus, their effects are described as *derived* or *emergent*, arising at least in part from the relational history with the stimuli in question. This ability to override direct acting relational contingencies and produce novel relations has allowed behavior analysts to talk in terms of the “verbal construction” of knowledge when applied to complex behavior (Hayes, 2004). In so doing, it has opened up exciting new vistas for the functional analysis of complex human behavior.

### **Clinical relevance of arbitrarily applicable relational responding research**

The research described here shares one important feature: a reliance on procedures which incorporate AARR. As we outlined earlier, it is AARR which gives human behavior, in all its glory, and gory, its symbolic splendor. It is AARR which allows humans to transcend the mundane and experience joy (at least fleetingly). It is AARR which enables scientists to predict and control, with ever increasing accuracy, the journeys of rockets through space and time. Indeed, it is AARR which gives us humans the experience of “time” (Hayes, 1992). And it is AARR which, at least partially, “traps” people in unhappy relationships, unfulfilling careers or in the safety of their homes. It is AARR which underlies the seemingly unique human tendency to spend one’s life worrying about the future, locked in a compulsive cycle of avoidance based coping with potential danger lurking everywhere.

The clinical relevance of AARR research is particularly evident in the transfer or transformation of stimulus functions (Dymond & Rehfeldt, 2000; Dymond & Roche, 2009); here, if one member of a relation is discriminative for a psychological response or function, like directly learned Pavlovian fear for instance, then other members of the relation will also

change and come to elicit fear, without further conditioning. That is, if A is the same as B and B is the same as C, and C is paired with an aversive or unpleasant event, then presentations of A will also come to actualize fear, in the absence of the event. According to this perspective, AARR may offer a functional account of the etiology and maintenance of some of the complex and seemingly puzzling patterns of non-perceptually evoked fear and avoidance responses often seen in anxiety disorders and which arise in the absence of an identifiable conditioning history (Coelho & Purkis, 2009; Dymond & Roche, 2009). For instance, there are occasions in the life of anxious individuals where merely thinking about or hearing a word related to the object of their fear or avoidance, such as “panic attack” or even other, seemingly unrelated words such as “anxiety” or “trapped”, pictures of someone suffering a panic attack or places where panic attacks may occur, can all come to suddenly occasion fear and avoidance (Foa & Kozak, 1986). In effect, while the original fearful CS may have been directly conditioned, emergent fear and avoidance responses readily occur along an arbitrary/relational dimension encompassing a plethora of indirectly related stimuli.

Interestingly, equivalence relations are just one type of AARR; relations such as opposition, comparison, temporal order (before/after), and spatial based forms of responding (e.g., May, Stewart, Baez, Freegard, & Dymond, 2017) illustrate the use of the terms of AARR to convey the operant basis of the effects described. The generativity of the effects obtained is increased when relations other than equivalence are involved; for instance, Dymond and Barnes (1995) demonstrated the emergence of three distinct patterns of behavior after training with one pattern, while Dougher et al. (2007) showed increases and decreases in skin conductance when transformation of functions occurred through derived comparative (more than/less than) relations. So, if A is more than B and B is more than C, and B is paired with shock, then presentations of A will evoke more fear than B, and C will evoke less fear than B in the absence of further shock.

In this article, we review the burgeoning AARR research on two aspects of human emotional suffering: fear and avoidance. Our aim is to describe the existing literature (as of mid-2017) and explore potential overlap and future synergies with the separate yet related domains of behavior analysis (Dymond & Roche, 2009), experimental psychopathology (Boddez, Davey, & Vervliet, 2016) and neuroscience (Schlund et al., 2016) for the empirical analysis of fear and avoidance in humans. We consider, for the first time, the validity of this work and address some of the therapeutic implications of a contemporary, behavior-analytic approach to AARR before describing several challenges and opportunities that may lie ahead.

### **Anxiety disorders and arbitrarily applicable relational responding**

Anxiety disorders have substantial healthcare costs. In Europe alone, where anxiety disorders have lifetime prevalence rates of 21%, and affect 41 million people, the estimated public health care costs exceed €41 billion (Wittchen, Jonsson, & Olesen, 2005). Despite extensive research into the underlying neurobiological mechanisms of anxiety and fear, “the causes of anxiety disorders remain largely unknown. This lack of certainty hinders accurate diagnosis, the prediction of prognosis, and the development of refined treatment approaches.” (Baldwin et al., 2010, p. 428).

Anxiety and related disorders appear irrational because they involve excessive fear and avoidance of seemingly arbitrary events and situations like harmless animals, invisible germs and enclosed spaces. Anxious individuals may have never directly experienced aversive consequences with these events in the past, yet the events come to evoke threat related responses, hyper-arousal, public and private sensations, and motivate high rate avoidance behavior (Guinther & Dougher, 2015). To some extent, then, part of the debilitating nature of anxiety-related disorders stems from the remote (i.e., derived) connection with, or total absence of, a prior history between humans and events that are

subsequently feared or avoided (Friman, Hayes, & Wilson, 1998). For the anxious individual, everything, previously experienced or not, is related to anxiety.

Accounting for the emission of fear and avoidance responses under such conditions requires more than an appeal to the physical, nonarbitrary controlling features of the environment (e.g., stimulus generalization; Dunsmoor & Paz, 2014; Dymond, Dunsmoor, Vervliet, Roche, & Hermans, 2015). That is, while novel, perceptually similar stimuli readily evoke conditioned responses, it is not always possible or indeed practical to identify physical similarities between feared objects, if they exist at all (Hermans & Baeyens, 2013; Hermans, Baeyens & Vervliet, 2013). Perceptual similarity is not the only way in which fear and avoidance may be indirectly acquired. Arbitrarily related public and private events that lack any physical resemblance to a prior conditioned stimulus can, and do, evoke and maintain excessive, maladaptive fear and avoidance. Indeed, “the powerful human capacity for abstract representation creates special problems in translating the rules of stimulus generalization worked out in animals. Fear and avoidance spread in animals from one context to another based on simple sensory cues. In humans, this spread may be on the basis of complex feelings” (Marks, 1987, p. 234). Complex feelings, or public and private responses to public and private events, are arbitrary or symbolic in nature (De Houwer, Barnes-Holmes, & Barnes-Holmes, 2017; Friman et al., 1998). That is, complex feelings may be discriminative for a range of maladaptive behavior in ways lacking a formal, physical correspondence between the private event and the referent. To this extent then, private events like thoughts and feelings are arbitrary responses, capable of modification via social whim (Hayes et al., 2001). Accordingly, the burgeoning literature on nonarbitrary (perceptual) generalization will not be reviewed here (Dunsmoor & Paz, 2014; Dymond et al., 2015; Hermans & Baeyens, 2013; Hermans et al., 2013).

**Arbitrarily applicable relational responding research on fear learning and extinction**



The differential fear-conditioning paradigm is widely adopted in preclinical, clinical and experimental psychopathology research on emotional behavior such as fear (Boddez, Baeyens, Hermans, & Beckers, 2014; LeDoux, 2014; Lonsdorf et al., 2017). The paradigm involves an aversive unconditioned stimulus (US), such as electric shock, being paired with a neutral conditional stimulus (CS+), while another stimulus (CS-) is paired with the absence of the US. Presentations of the CS+, but not the CS- then come to reliably evoke conditioned responses (CRs) such as freezing/orienting responses and increased physiological arousal (i.e., changes in skin conductance, heart rate, respiration, pupil dilation) in the absence of the US. Extinction is usually achieved by repeated unpaired presentations of CS+, leading to diminution of learned fear. The clinical implications of fear-conditioning have been extensively studied for decades and enormous contributions made to the basic and applied analysis of emotional systems in humans and nonhumans (Bouton, Mineka, & Barlow, 2001; Craske, Hermans, & Vansteenwegen, 2006; LeDoux, 2014). As a laboratory paradigm, fear conditioning and extinction procedures form a central part of the studies reviewed here.

Early laboratory-based demonstration work on AARR and fear tended to employ repeated acquisition, small-*n* research designs and electrodermal (skin conductance) measures of learning and transfer (Dougher, Augustson, Markham, Greenway, & Wulfert, 1994). For instance, Dougher et al. (1994) first trained and tested two, four-member equivalence relations consisting entirely of arbitrary line drawings (A1-B1-C1-D1 and A2-B2-C2-D2) in eight healthy, female participants. Next, using Pavlovian fear conditioning procedures, one member of the first relation (B1) was established as a CS+ by pairing it with shock, and a member of the second relation (B2) was established as a CS-. Transfer of fear-eliciting functions was then tested with a fixed order (counterbalanced across participants) of presentations of the CS+, which continued to be occasionally followed by shock, CS-, which continued to not be followed by shock, and all other C and D stimuli from the two relations,

which were presented in the absence of shock. Skin conductance was the primary outcome measure used to infer conditioning and transfer, and per this rubric, 6 out of 8 participants showed evidence for the transfer of Pavlovian fear-eliciting functions through equivalence relations. In a second experiment, Dougher et al. (1994) replicated these findings with a further four participants and extended the demonstration to the transfer of extinction through equivalence. In that experiment, fear was first conditioned to several members of one relation (B1, C1 and D1), extinguished for one (B1) and then tested for transfer of extinction to the remaining members of the relation (C1 and D1). Conditioning and transfer of fear was then (re)tested. All four participants, and none of the four control participants who did not receive the relational training and testing phases, demonstrated transfer of extinction through equivalence relations. Further studies replicated and extended this basic effect (Augustson, Dougher, & Markham, 2000; Markham, Dougher, & Augustson, 2002; Markham & Markham, 2002; see also, Smyth, Barnes-Holmes and Forsyth, 2006).

Twenty years following Dougher et al.'s initial demonstration, Vervoort, Vervliet, Bennett, and Baeyens (2014) repeated the study using a between-groups experimental design, large sample size, and inferential statistical analysis. As predicted, fear, which was now also inferred from trial by trial US expectancy ratings, transferred to related (C1, D1) but not unrelated stimuli (C2, D2). Vervoort et al. also found that extinction of fear also transferred to related stimuli, while extinguishing fear of a derived stimulus (e.g., C1) did not reduce fear of the original CS+.

Rodriguez-Valverde, Luciano, and Barnes-Holmes (2009) also addressed several methodological factors in the Dougher et al. study, such as the absence of a control group in the first experiment, the presentation order of the crucial test probes, some of the conditioning parameters (i.e., CS-US interval), and skin conductance response quantification. No evidence was found for the transfer of fear eliciting functions under these conditions. In another

experiment, a large sample of participants were exposed to revised conditioning parameters and response quantification and evidence for transfer was found in most of the participants. As Rodriguez-Valverde et al. opined, the relevance of these findings for experimental psychopathology research is enhanced when more conventional research and analysis paradigms are used.

In the only study to our knowledge to include autonomic measures other than skin conductance, Medina, Valverde, and Lopez (2016) found no evidence for conditioned or derived transfer of fear-potentiated eye-blink startle through equivalence relations. Eye-blink startle, which is measured via electromyography electrodes applied to the orbicularis oculi muscle, is a well validated measure in experimental psychopathology research (Grillon, 2008), and despite these initial negative findings may still hold potential for future AARR research on fear.

The separate and combined effects of AARR and nonarbitrary, perceptual similarity was recently examined by Bennett, Vervoort, Boddez, Hermans and Baeyens (2015). Fear (over-)generalization is thought to involve the transfer of fear-eliciting functions to innocuous stimuli by virtue of their *similarity* to other threat-relevant stimuli. This similarity can be perceptual when stimuli reside along a single physical continuum (Vervliet, Vansteenwegen & Eelen, 2004). It might also be arbitrary when physically dissimilar stimuli share a pre-experimentally learned conceptual property, such as category memberships (e.g., animals versus flowers; Dunsmoor & Murphy, 2015); that is, conceptual forms of generalization are in effect demonstrations of arbitrary generalization via predefined conceptual categories. Moreover, it is likely that events in conditioning episodes involve *both* perceptual and conceptual details. During a road traffic accident, for instance, a victim might attend to the physical form of an approaching car (e.g., its burgundy red color) as well as more arbitrary or symbolic details (e.g., a 'Ford'-made car). Afterwards, fear could generalize to novel stimuli

along both a perceptual dimension (e.g., a burgundy red wine) and arbitrary dimension (e.g., a Ford radio advertisement or a map showing where the accident occurred).

These forms of generalization are rarely studied together. This disparity is potentially problematic as researchers may perpetuate a false dichotomy wherein fear generalization is regarded as either perceptual or non-perceptual (Moors, 2014), which could impact the external validity of this research. With this in mind, Bennett and colleagues (2015) examined whether the perceptual and conceptual features of stimuli concurrently facilitate fear generalization. Two, three-member stimulus equivalences classes were established using a matching-to-sample task and a member of one class (CS+) was then associated with an aversive US. The generalization of fear to (i) other category members and (ii) perceptual variants of these members was examined. As predicted, both conceptual and perceptual similarity facilitated the transfer of fear and avoidance. Previously neutral members of the aversive class, as well as their perceptively similar variants, elicited heightened threat expectancy ratings and avoidance behavior. These findings, which are particularly noteworthy given that most research on fear generalization has been conducted on perceptual-based generalization (Dunsmoor & Paz, 2015; Dymond et al., 2015), highlight how conceptual (i.e., arbitrary) and perceptual stimulus relations can intensify the spread of fear (Dunsmoor, Kroes, Braren, & Phelps, 2017). In real-life, this combination might create a nexus of events that elicit intense fear and debilitating avoidance.

Further extensions of AARR research on fear and extinction learning may be found in the literature concerned with understanding chronic pain. Patients with chronic pain often report pain-related fear of engaging in specific movements that have never featured in painful episodes (Meulders & Vlaeyen, 2013) and some researchers have appealed to stimulus generalization to understand this outcome. If one arm movement is paired with a painful outcome, for example, then proprioceptively similar arm movements are found to partially

elicit fear (Meulders, Vandebroek, Vervliet, & Vlaeyen, 2013). Bennett, Meulders, Baeyens, and Vlaeyen (2015) recently suggested that AARR might also afford problematic over-generalization of pain-related fear in chronic pain disorders. In this study, stimulus equivalence classes were established in which nonsense words and joystick arm movements were equivalent. Using a differential fear conditioning paradigm, nonsense words were then associated with a pain-US. During the critical test phase, joystick arm movements from within the same stimulus equivalence category elicited pain-related fear, despite the absence of the pain-US on these trials. This finding highlights how pain-related fear might emerge in the absence of a painful episode.

As a secondary objective, Bennett, Meulders et al. (2015) also compared two fear acquisition methods - direct CS-US conditioning or verbal instructions about the CS-US relation, respectively. Using a between-groups design, nonsense words were associated with a pain-US through either direct experience or through verbal information. Both direct learning and verbally instructed pathways evoked the AARR-based generalization of pain-related fear to joystick arm movements. In addition, direct experience with CS-US presentations resulted in stronger acquisition and generalization effects. Relative to the verbal information group, participants who experienced the CS-US pairings, firstly, reported the aversively conditioned stimulus as being more unpleasant and, secondly, the conceptually similar movements as being more fear-relevant and unpleasant. This may be valuable information when considering the etiology of anxious suffering and coping with chronic pain.

### **Beyond Equivalence and Fear: Multiple stimulus relations**

To date, the majority of AARR research on fear has been concerned with equivalence relations. However, one study that employed stimulus relations other than equivalence is Dougher, Hamilton, Fink, and Harrington (2007), who examined the transformation of fear-eliciting functions within a 3-member comparative (more than/less than) relational network.

During this study, participants were first exposed to a relational training phase in which they were presented with one of three arbitrary sample stimuli (A, B and C) in the top portion of the computer screen, and three comparison stimuli in the bottom portion of the screen. The comparison stimuli were physically similar but differed in terms of size (e.g., small, medium and large). The purpose of this phase was to train participants to select the smallest comparison in the presence of stimulus A, the medium comparison in the presence of stimulus B, and the largest comparison in the presence of stimulus C. Once participants met criterion at testing, they were exposed to a bar press training and test phase. They were initially trained to press the spacebar at a steady rate to the B stimulus before being exposed to a test phase involving trials on which either the A or C stimulus was presented by itself. No instructions were presented on these trials, and the dependent variable was the rate at which participants pressed the bar in the presence of the A and C stimuli. Results demonstrated that participants pressed slower to stimulus A and faster to stimulus C, than they did to B. During a subsequent phase of the experiment, participants were exposed to Pavlovian conditioning with stimulus B, and testing with stimuli A and C. Thus, stimulus B was paired with a mild shock, and changes in skin conductance were employed as the dependent variable. The researchers found that 6 out of 8 participants demonstrated smaller skin conductance changes to stimulus A, and larger changes to stimulus C, than to stimulus B. To our knowledge, Dougher et al. (2007) remains the only study on multiple stimulus relations and fear-eliciting functions.

#### **Arbitrarily applicable relational responding and avoidance**

Avoidance behavior is a central diagnostic factor and prominent clinical feature in many anxiety disorders, such as generalized anxiety disorder and specific phobia (American Psychological Association, 2013; Craske, Rauch, Ursano, Prenoveau, Pine, & Zinbargh, 2009). Once established, avoidance is difficult to extinguish and may become the default means of coping with real and perceived danger and the situations associated with them.

Thus, avoidance can often become maladaptive, leading to diminished opportunities to engage in approach related behavior and impaired daily functioning with an elevated risk of psychopathology and comorbid disorders.

In the lab, several approaches to the study of avoidance have been developed (Dymond & Roche, 2009; Higgins & Morris, 1984; Kryptos et al., 2014; LeDoux, Moscarello, Sears, & Campese, 2016). In classic avoidance procedures, a discrete avoidance response in the presence of a warning stimulus prevents or minimizes contact with an aversive event. Classic avoidance may be active, in which an overt response by the organism is required, or may be passive, in which the aversive events are avoided by withholding responses. Arbitrarily applicable relational responding research on avoidance has tended to employ active, differentiated avoidance tasks in which avoidance responses performed in the presence of a fear conditioned cue (CS+) avoid scheduled aversive events, while avoidance in the presence of a cue paired with the absence of the US (CS-) resulted in no programmed consequences. In what follows, we describe the findings of this AARR work, which has mainly been investigated with active avoidance.

In the first study to investigate derived avoidance, Augustson and Dougher (1997) trained and tested participants for the formation of two, four-member stimulus equivalence relations (A1-B1-C1-D1 and A2-B2-C2-D2) and then established B1 as CS+ and B2 as CS-, respectively. During a subsequent avoidance-training phase, shock could be avoided by completing a fixed-ratio (FR) 20-response requirement in the presence of B1, while shock was never scheduled to follow presentations of B2. Derived avoidance was then tested, in the absence of shock, with presentations of the indirectly related stimuli (C1, D1, C2, and D2) that were not present during avoidance training. Findings showed that all eight participants emitted the avoidance response to C1 and D1 but not to C2 and D2, indicating that the avoidance functions trained to B1 transferred to C1 and D1 without further training.

Since the seminal study by Augustson and Dougher (1997), others have replicated and extended this basic effect. Dymond, Schlund, Roche, Whelan, Richards, and Davies (2011) employed a larger sample size in an avoidance paradigm involving aversive images and sounds as USs, while Dymond, Schlund, Roche, De Houwer, and Freegard (2012) showed that levels of derived avoidance resemble those seen when avoidance is acquired through verbal instructions (Rachman, 1977).

Boyle, Roche, Dymond, and Hermans (2016) examined whether natural language categories facilitate the generalization of fear and conditioning from conditioned cues to their synonyms. In that study, participants were exposed to a fear conditioning procedure in which a commonly used English word (e.g., *broth*) was paired with a brief electric shock in a classical conditioning paradigm. Another English word (e.g., *assist*) was designated as the CS- and was never followed by shock. Subsequently, avoidance responses were conditioned by providing an avoidance response option in the presence of the conditioned fear stimuli. In the generalization test phase, synonyms of the CS+ (e.g., *soup*) and CS- (e.g., *help*) were presented in extinction. Levels of avoidance, skin conductance and post-hoc expectancy of shock ratings were significantly higher for the CS+ than the CS-, but more interestingly, they were also higher for the synonym of the CS+ compared to a synonym of the CS-. This was the first study to demonstrate the generalization of both fear and avoidance across naturalistic semantic relations and has implications for the way in which such relations are modelled in the AARR laboratory.

According to relational frame theory (RFT; Dymond & Roche, 2013), excessive avoidance occurs in the presence of a wide range of stimuli and situations based on the actual and, more often than not, inferred presence of the aversive event. If AARR is responsible for the excessive levels of maladaptive avoidance and other responses to threat often seen in phobias and anxiety related disorders, then it follows that groups that differ on clinically



relevant trait variables may show different levels of derived avoidance. To our knowledge, only one study to date has recruited participants with sub-clinical individual differences (Dymond, Schlund, Roche, & Whelan, 2014). In this study, participants with and without fear of spiders first learned two stimulus equivalence classes ( $A1=B1=C1$  and  $A2=B2=C2$ ). Images of spiders (US) were then paired with a member of one stimulus equivalence class (B1) but not the other (B2). Students with elevated fear of spiders subsequently demonstrated heightened avoidance and US expectancy ratings, relative to those students with little fear of spiders. These findings demonstrate that fearful participants show derived avoidance at higher levels than nonfearful participants, which supports the assertion made by RFT that AARR is crucial in the acquisition and maintenance of clinical behavior. It is possible, therefore, that “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” and that “spiders may also have had different functions for the phobic group (e.g., fear) compared to controls (e.g., ambivalence, curiosity, disgust)” (Stewart, Stewart, & Hughes, 2016, p. 241). Clearly, further work is needed to test the role of AARR in clinically relevant avoidance.

Other AARR work has been concerned with extinction or reduction of avoidance. Luciano et al. (2013) explored context-specific effects of exposure on derived avoidance and found that CS extinction in one context had a minimal effect on avoidance responding in another context. Luciano et al. (2014) found that an intervention drawn from Acceptance and Commitment Therapy (ACT) reduced levels of derived avoidance compared with a control intervention, while Garcia-Guerro, Dickins and Dickins (2014) found that extinction of derived avoidance was facilitated by an instructional prompt, but that considerable individual differences were observed in the gradient of responding during extinction. The small sample sizes and multi-phase nature of these studies suggest that further empirical attention is warranted on extinction of derived avoidance.

**Beyond Equivalence and Avoidance: Multiple stimulus relations**

Unlike fear, a growing number of AARR studies on avoidance have employed relations other than equivalence. For example, Dymond, Roche, Forsyth, Whelan, and Rhoden (2007, 2008) showed that derived avoidance may be transformed in accordance with relations of “sameness” (i.e., equivalence) and “opposition”. After training to establish two abstract shapes as contextual cues for same (S) and opposite (O), respectively, the cues were presented with arbitrary nonsense stimuli, and participants were taught the following relations: S-A1-B1, S-A1-C1, O-A1-B2 and O-A1-C2. These relations lead to the following eight untrained relations: B1-C1 are same, C1-B1 are same, B2-C2 are same, C2-B2 are same, B1-C2 are opposite, C2-B1 are opposite, B2-C1 are opposite, and C1-B2 are opposite. It was predicted that participants would (a) choose C1 given B1 and choose B1 given C1 in the presence of S (C1 and B1 are both the same as A1 and therefore the same as each other); (b) choose C2 given B2 and choose B2 given C2 in the presence of S (C2 and B2 are both opposite to A1 and therefore the same as each other); (c) choose C2 given B1 and choose B1 given C2 in the presence of O (C2 is opposite to A1, and B1 is the same as A1, and therefore C2 is opposite of B1); (d) choose C1 given B2 and choose B2 given C1 in the presence of O (C1 is the same as A1, and B2 is opposite to A1, and therefore C1 is opposite to B2).

Participants were then exposed to a signaled avoidance task, during which responding in the presence of stimulus B1 cancelled a scheduled US presentation. Another stimulus, B2, was never followed by the US. Then, participants were tested with presentations of C1 and C2, in the absence of the US. Findings showed that consistent avoidance responses were made in the presence of C1 but not C2 (because C1 is the same as B1, whereas C2 is the opposite), thus demonstrating derived avoidance in accordance with complex relational networks of same/opposite (Gannon, Roche, Kanter, Forsyth, & Linehan, 2011; Roche, Kanter, Brown, Dymond, & Fogarty, 2008).

In an extension of this work, Bennett, Hermans, Dymond, Vervoort, and Baeyens, (2014) tested whether the stimulus established as opposite to the CS+ (C2) evoked minimal avoidance because it participated in a derived sameness relation with a safe CS rather than participating in a relation of derived opposition with an aversive CS (shock). Bennett et al. therefore conducted a systematic replication of Dymond et al.'s (2007, 2008) procedures to rule out sameness with a non-aversive CS as an explanation for low levels of avoidance to the opposite stimulus from the network. Testing this relational interpretation of derived avoidance involved conducting a non-differential fear conditioning phase in which only the putative CS+ (X1) was presented; that is, stimuli related by opposition were never paired with the absence of the aversive stimulus. During the critical test for derived avoidance, groups of participants were presented with X1, X2, a novel stimulus and either Y1 or Y2 (as both were in derived opposition with X1). Findings indicated that the derived same stimulus, X2, evoked a higher proportion of avoidance responding than either of the derived opposition stimuli, Y1 or Y2, or the novel stimulus. Self-report measures of US expectancy and stimulus valence were in line with these findings. Bennett et al.'s findings are significant because they demonstrate instances where derived opposite relations suppress derived avoidance relative to derived same relations or a neutral stimulus.

### **Check the validity: Applying validity criteria tests to AARR research on fear and avoidance**

To make the case that AARR explains, at least partially, real-world cases of complex derived fear and avoidance (Dymond, Roche & Bennett, 2013), it is critical to demonstrate the ecological validity of the relevant laboratory research. Vervliet and Raes (2013) described four criteria to assess whether a laboratory procedure ("the model") has ecological validity. These include (i) *face validity*, (ii) *diagnostic validity*, (iii) *predictive validity* and (iv) *construct validity*. Here, we apply these criteria to gauge the ecological validity of AARR

procedures with respect to fear and avoidance in human anxiety (see also Scheveneels, Boddez, Vervliet, and Hermans, 2017).

*Face validity.* Face validity is “the degree of phenomenological similarity between the behavior in the model and the symptoms of the disorder” (Vervliet & Raes, 2013, p.2241). With respect to the present literature, face validity pertains to the extent to which individuals with anxiety disorders demonstrate fear and avoidance to symbolically related events. In fact, instances of symbolic or derived fear are common across a range of human anxiety disorders. In OCD, for example, a catalog of items that are physically dissimilar (e.g., washing powder, ammonia, cleaning gels, aerosols) but conceptually alike (i.e., corrosive household agents) trigger heightened fear and repetitive coping behavior (Hermans et al., 2013). A person with specific blood and injection phobia, for example, might also dread arbitrarily related stimuli like white hospital coats or the scent of a hospital (Dunsmoor, Martin, & LaBar, 2012). Such cases can be functionally interpreted to involve stimuli which evoke conditioned (fear) responses after an experience with physically dissimilar stimuli (Boddez, Bennett, van Esch, & Beckers, 2016).

The gold-standard AARR protocol usually consists of the following: one member of a particular derived stimulus network (e.g., equivalence) is paired with an aversive outcome such that other members are treated as if they were previously associated with said outcome (Dymond et al., 2011). Moreover, with relational networks of opposition, if one stimulus is associated with an aversive outcome, then participants respond to stimuli related via derived opposition relations *as if* they were negatively correlated with that outcome (Bennett et al., 2015c; Dymond et al., 2008). It appears then that both the experimental protocol and real-world examples involve the elicitation of heightened fear in response novel stimuli, following the pairing of a physically dissimilar stimulus with an aversive outcome (Dymond et al., 2011). There is, therefore, assumed functional overlap between the experimental procedures

used in AARR research and real-world cases of derived fear and avoidance. Thus, AARR research on fear and avoidance arguably possesses face validity.

*Diagnostic validity.* Diagnostic validity refers to findings obtained with clinical or sub-clinical populations showing that “the behaviors differ from healthy individuals (in intensity or frequency)” and as a result, “the model can be used as a diagnostic marker” (Vervliet & Raes, 2013, p. 2242). With respect to AARR and anxiety disorders, diagnostic validity requires that anxious participants show more derived fear reactions or transformation of anxiety inducing functions than non-anxious participants. As mentioned previously, most AARR research studies use non-anxious student participants (cf. Leslie, Tierney, Robinson, Keenan, Watt, & Barnes, 1993) and only one experimental study has used sub-clinically spider-fearful participants (Dymond et al., 2014). Dymond et al. showed that high spider fearful individuals exhibited greater levels of derived avoidance and met avoidance learning criteria quicker than low spider fearful individuals. While this is an encouraging first step, it remains unclear whether a heightened transfer of emotionally-relevant behavior is specific to anxiety disorders or is instead a general characteristic of a large range of clinical disorders.

Despite its impressive volume, AARR research is skewed towards proof-of-principle studies that merely demonstrate derived fear-relevant behavior. There are considerably fewer published attempts to extend beyond these foundational designs. Consequently, it remains unclear as to whether AARR predicts, or even relates to, abnormal fear derivation processes in human anxiety. That is, the diagnostic validity of this research remains unknown (Vervliet & Raes, 2013).

Considerably more research is required to learn how AARR might contribute to the spread of fear and avoidance in human anxiety. Cross sectional research studies that recruit participants with different levels/types of anxious symptomatology are required to elucidate the relationship between AARR and maladaptive fear (Lissek, 2012). Longitudinal research

studies are also required to determine whether AARR plays a causal role in the emergence of maladaptive fear. For instance, a study could investigate whether the derived transfer of fear predicts subsequent reports of arbitrary, derived fear as well as other anxious symptoms (Lenaert, Boddez, Griffith, Vervliet, Schruers, & Hermans, 2014).

*Predictive Validity.* “Predictive validity means that performance in the model predicts performance in the disorder” (Vervliet & Raes, 2013, p. 2241). One way of evaluating the predictive validity involves testing whether variables which influence performance in the experimental protocol also influence clinical treatment of the real-world phenomenon. This approach has received considerable support in the experimental psychopathology literature but less so in AARR research on fear and avoidance. Fear extinction, for instance, which is the repeated exposure of fear conditioned stimuli in the absence of an aversive outcome, is an effective fear reduction strategy (Craske, Treanor, Conway, Zbozinek, & Vervliet, 2014; Scheveneels et al., 2017). Fear extinction has also been shown to partially diminish the transfer of conditioned arousal and threat-expectancy ratings (Dougher et al., 1994; Vervoort et al., 2014). However, to fully establish predictive validity of AARR research on fear and avoidance, it will be necessary to demonstrate that those individuals who respond more robustly on tests for derived transfer of fear/avoidance also are more susceptible to developing anxiety disorders. At present, the only evidence we have for predictive validity is indirect, but in line with the prediction.

There is therefore a limited number of studies to suggest the predictive validity of AARR research on fear and avoidance. Psychotherapeutic interventions that partially attenuate real-world fear do however also diminish the transfer of fear-related behavioral functions in the laboratory. Ultimately, further research is needed to develop and optimize evidence-based fear reduction techniques involving AARR processes (Craske et al., 2014).

*Construct validity.* Construct validity refers to “the disease relevance of the methods by which the model is constructed, with a focus on recreating the etiological process in the model” (Vervliet & Raes, 2013, p.2242). A strongly developed measure of construct validity necessitates an “elaborated (etiological) theory of the disorder and of the model, and theoretical reasons to assume that the process in the model parallels the clinical process of interest.” (Vervliet & Raes, 2013, p.2242). Thus, a theoretical framework or underlying rationale is needed to explain why experimental procedures employed in AARR research mirror the psychological phenomenon under study. In this case, in what way is AARR implicated in derived fear and avoidance?

A common assertion by AARR researchers is that anxious symptomatology (e.g., maladaptive fear and avoidance responses) are rooted in maladaptive verbal behavior. As a functional account of verbal behavior, AARR is thought to be a problematic source of behavioral control that, in humans, can tend to dominate over actual contingency-governed experiences (Hayes, Levin, Plumb-Villardaga, Villatte, & Pistorello, 2013; Stewart et al., 2016; Villardaga, Hayes, Levin, & Muto, 2009). Thus, a range of verbal stimuli come to evoke anticipatory fear and avoidance despite never featuring in an aversive context (Dymond et al., 2013; Friman et al., 1998). While this proposed rationale is plausible, there is an absence of direct evidence. There is, firstly, a limited amount of information on whether the transfer or transformation of fear-relevant functions is heightened in anxious populations (i.e., poor diagnostic validity; cf. Dymond et al., 2014). Second, there is no clear, accepted explanation about why the events that are arbitrarily related in AARR might lead to maladaptive stimulus control and behavior in the first place.

To some within behavior analysis, the work reviewed here amounts to “laboratory purported-analogs” (Todd Risley, quoted in Rutherford, 2009, p.60). According to this viewpoint, there is little obvious value in translating real-world, clinical problems to a

laboratory analog task and attempting laboratory based demonstrations of the putative behavioral processes involved because, while they are intended to “shed light on socially important problems”, they do without establishing any “socially-valued behavior change” (Critchfield & Reed, 2017). However, the modelling of clinically significant phenomena, such as fear and avoidance responses, under controlled laboratory conditions is a widely-adopted practice in clinical psychology and experimental psychopathology research (Scheveneels et al., 2017). Indeed, this approach has served experimental psychopathology extremely well and has, for instance, lead to the direct translation of lab-based findings to clinical intervention aimed at augmenting exposure therapy for anxiety disorders (e.g., Craske et al., 2014). Moreover, the validity of experimental psychopathology work is well established (Vervliet & Raes, 2013) and researchers within this approach are continually refining and improving their laboratory paradigms and the range of measures employed (e.g., Meulders et al., 2013). The same cannot be said about the behavior analysis of fear and avoidance, apart perhaps from the advances made in behaviorally-informed therapies such as ACT (Hayes et al., 2014; Levin & Villate, 2016). Thus, while the nature and status of laboratory-based treatment research for revealing clinically relevant processes is widely accepted in domains outside of behavior analysis, we maintain that there is much to recommend such an approach for the behavior-analytic study of emotion.

To some, it is likely that findings obtained from the laboratory-based research we have reviewed here are deemed irrelevant because, as Critchfield and Reed (2017) state, “the setting is artificial, the participants may be selected at least partially for convenience, and/or the behavior under study bears only partial similarity to that seen in everyday circumstances. Nevertheless, the relevant experiments may illuminate mechanisms that matter in everyday settings and clinical interventions”. We endorse this optimistic conclusion and call for further laboratory based treatment studies on AARR in fear and avoidance; there is a need to study



these processes under controlled conditions to eliminate potential confounds and because demonstrating these processes in the laboratory is a necessary first step in the broader research process aimed at translational extension and clinical application. Questions for future research include, what is the learning history necessary for heightened transfer of fear-relevant behavior and how prevalent is this learning history in the lives of anxious individuals? It seems, therefore, that the underlying connection between AARR and anxious symptomatology is relatively underdeveloped.

Fortunately, these criteria of validity are not independent; as the research progress and the diagnostic validity becomes clearer, for instance, we will be in a better position to propose theoretical reasons for why our experimental models reflect AARR process in maladaptive fear and avoidance.

### **Therapeutic Implications of AARR Research on Fear and Avoidance**

Research into AARR and the transformation of fear and avoidance response functions as a core process in the spread of fear and avoidance has obvious therapeutic implications (McEntegart, Barnes-Holmes, Hussey, & Barnes-Holmes, 2015; Villate, Vilatte, & Hayes, 2017). Not only does research of the kind reviewed here speak to the etiology of a range of anxiety conditions, it also provides suggestions as to how well-established and emerging treatments for anxiety disorders may work. For example, cognitive restructuring is widely used in the treatment of fear and avoidance related conditions (Clark & Beck, 2012). This method focuses on testing the often-false beliefs of clients regarding the need for and benefits of responding fearfully in problematic situations. Clients are invited to reality-test their own beliefs and to consider contradictions between them. By helping clients to see the erroneously catastrophic nature of their fears and the unrealistic nature of the outcome of not avoiding feared situations, behavior can begin to change in line with “reality” (i.e., operating contingencies).

Within clinical behavior analysis, a parallel but differently conceived treatment technique known as cognitive defusion (Blackledge, Moran & Ellis, 2009; Hayes, Strosahl & Wilson, 2012) is more likely to be employed in the treatment of excessive fear and avoidance. Unlike cognitive restructuring, the aim of defusion is not to create new or competing verbal relations in which aversive stimuli are related. Rather, defusion attempts to reduce the impact of already-established verbal relations on subsequent behavior without in any way altering the form or frequency of those verbal relations (Blackledge, 2007; 2015; Hayes, Luoma, Bond, Masuda & Lillis, 2006). This is achieved via a series of exercises that alter the functional context of problematic verbal relations. For example, in a word-repetition exercise, clients are encouraged to repeat a problematic word aloud to observe how the various poorly discriminated functions of the word, such as its sound, can often become more salient, and the highly dominant ones, such as the fear it produces, less salient with repetition. In variations of this exercise the client is asked to repeat emotionally challenging statements in a “silly voice” (e.g., “I am afraid of dogs” spoken as a cartoon character) to observe how it changes the impact of the statement. To date, several studies have reported a positive impact of these types of functional context-changing (i.e., defusion) exercises on the believability of negative thoughts using a variety of different means of inducing those thoughts (e.g., Barrera, Szafranski, Ratcliff, Garnaat & Norton, 2015; Hooper & McHugh, 2013; Masuda, Hayes, Sackett & Twohig, 2004; Ritzert, Forsyth, Berghoff, Barnes-Holmes & Nicholson, 2015).

Other putatively different defusion exercises involve the client acting in ways that are incoherent with ongoing statements (e.g., repeating “I can’t walk”, while walking), to allow the client to discriminate the limits of their verbally based thoughts in controlling behavior (Tyndall, Papworth, Roche, & Bennett, 2017). Also, included under the umbrella of “defusion” techniques are more mindfulness-based exercises (e.g., seeing thoughts as leaves on a stream) that involve discriminating or noticing without evaluating or responding to

thought itself to create a sense of space or distance between one's self and one's thoughts. These methods in some way change the transformation of function process, while leaving the problematic verbal relations (e.g., "I am afraid of dogs") intact. However, the very variety of defusion techniques may be a problem for a basic process account of this approach to treatment. Specifically, it is not at all clear if the processes underlying these clinical exercises are the same and researchers have begun to question whether cognitive defusion is a process, a technique or an outcome (Barnes-Holmes, Hussey, McEntegart, Barnes-Holmes, & Foody, 2016; Dymond et al., 2013; Foody, Barnes-Holmes, Barnes-Holmes, & Luciano, 2013; Levin & Villatte, 2016; Lopez-Lopez & Luciano, 2017).

Defusion exercises may simply produce derived extinction via exposure to stimuli *in vivo* or indirectly via narrative activity (e.g., derived transformation of extinction functions). Derived extinction of fear and avoidance has now been demonstrated in several studies (see earlier sections) and is relatively well understood. Indeed, researchers have even investigated directional effects in the transfer of derived extinction. That is, one study (Roche et al., 2008) found no difference in the rate of extinction transfer within a derived relation when the originally conditioned discriminative stimulus ( $S_D$ ) was presented in extinction over a procedure in which a "derived  $S_D$ " for avoidance was presented in extinction. In contrast, however, a more well-controlled study (Vervoort et al., 2014), discussed earlier, found that exposure to the conditioned  $S_D$  for avoidance does lead to more reliable transformation of extinction to related stimuli, than *vice versa*. This suggests that exposure should be more effective in producing reductions in fear and avoidance than "talk therapy", which often targets only the symbolically related stimuli (i.e., via narrative activity). This is precisely the type of process level research that is needed to disentangle potential core process at work in cognitive defusion.

One obvious and pressing question regarding the process of defusion (or derived

extinction) relates to the limits imposed on efforts to induce it by the relatedness of the stimulus presented in extinction (i.e., in an exposure analog) and the original conditioned fear or avoidance stimulus. Previous research has shown that increasing relational distance between related stimuli weakens the derived relatedness of stimuli and the transfer of functions (Moss-Lourenco & Fields, 2011). Thus, we might expect that talk therapies that target stimuli related only remotely to the original conditioned stimuli may have limited effects in terms of the derived transfer of extinction. While this relatedness effect has not been studied specifically in relation to relational frames, several studies have now shown complex fear and avoidance function transformation in accordance with comparative (e.g., Dougher, et al., 2007) and same/opposite relations (Bennett et al., 2015; Dymond et al., 2007). Accordingly, it is unclear how effectively extinction/exposure based treatments would work if the relations according to which fear and avoidance have been transformed in a real-life scenario are not known. For example, it is not yet known if stimuli that participate in a hierarchical verbal relation with a conditioned aversive US, can be targeted as effectively in an exposure treatment paradigm as a stimulus that participates in an equivalence relation with the same aversive stimulus. In effect, we do not yet know how extinction effects transform in accordance with extended equivalence relations or relations other than equivalence. Tackling these complex process issues and attempting to identify boundary conditions for derived fear and avoidance extinction effects are important next steps for experimental psychopathology research, behavior analysis and clinical behavior analysis.

### **Conclusions and Deriving the Future**

Beginning with the writings of William James, B. F. Skinner and others, behavior analysis has made substantial contributions to the study of human emotion (e.g., Forsyth & Eifert, 1996; Friman et al., 1998; Killeen & Jacobs, 2016; Lewon & Hayes, 2014; Skinner, 1945, 1953; Taylor & O'Reilly, 1997). Yet, while these works have engendered much debate

and occasional controversy (Lamal, 1998), a fully articulated case on the status of emotions in the science of human behavior and an accepted, coherent behavioral science of emotions remain lacking. Part of the reason why behavior analysis has done, and to some extent, still does, lag other approaches to the study of emotion is because it has only relatively recently developed ways of talking about and investigating, from a functional perspective, one of the key defining features of complex emotional behavior: the verbal, relational basis of human suffering. As we indicated above, it is now widely accepted that AARR provides the functional basis for a contemporary behavior-analytic understanding of emotion.

Fully grasping the potential of AARR for the behavior analysis of emotion will lead to new ways of studying emotion in the lab and to advances in therapeutic application (Dymond & Roche, 2009). It may also lead to new ways of talking (Vilatte et al., in press). That is, much of the recent work has adopted what are called middle-level terms to describe the behavioral processes under investigation. Middle-level terms are non-technical, theoretically distinct terms not developed from basic research. For instance, the burgeoning use of the terms, *symbolic generalization* (Bennett et al., 2015; Dymond et al., 2011, 2012; Dymond, Molet & Davies, 2017) to describe transfer or transformation of clinically relevant stimulus functions is intended to facilitate communication with other domains (Dymond et al., 2015). We should increase such efforts, as proponents of contextual behavioral science have maintained (e.g., Barnes-Holmes et al., 2016). Could *symbolic generalization* offer one such mid-level term, describing a functional outcome (the spread or generalization of fear or avoidance along an arbitrarily applicable stimulus dimension) in ways which other functional-analytic terms like AARR and transformation of stimulus functions cannot? We'd like to think so.

Further AARR applications in the neurosciences can aid in addressing questions about underlying neuropathology in various clinical disorders and basic research questions about

function in brain regions supporting derived threat and avoidance. We maintain there is much to be gained from this exchange because, unlike AARR approaches to studying fear and avoidance, the neural substrates of perceptually-based stimulus generalization are relatively well known in humans and nonhumans. Studies reveal that fear generalization engages similar neurocircuitry during both the acquisition (e.g., insula and anterior cingulate cortex) and regulation (e.g., ventromedial prefrontal cortex) of directly conditioned fear (e.g., Dunsmoor & Paz, 2014; Dunsmoor, Prince, Murty, & Kragel, 2011; Lissek, 2012; Lissek et al., 2014; Onat & Buchel, 2015). To date, however, no direct neurobehavioral investigations of AARR in fear and avoidance have been reported. Instead, a number of related human functional magnetic resonance imaging (fMRI) studies have been conducted that highlight a distributed brain network supporting AARR during tests of emergent performance on tasks such as transitive inference (TI; Acuna, Eliassen, Donoghue & Sanes, 2002; Królicki & Wróbel, 2011), acquired equivalence (Schlichting & Preston, 2015; Zeithamova, Dominick, & Preston, 2012), stimulus equivalence (Dickins et al., 2001; Ogawa, Yamazaki, Ueno, Cheng & Iriki, 2010; Schlund, Hoehn-Saric & Cataldo, 2007; Schlund, Cataldo & Hoehn-Saric, 2008), and arbitrarily applicable comparative relations (Hinton, Dymond, Von Hecker & Evans, 2010). There are several consistent regional patterns of brain activation across the published studies, which represent encouraging systematic replications. The activation maps in Figure 1 are drawn from these studies and highlight that AARR is supported by dorsal, ventral and medial portions of the prefrontal cortex, posterior parietal regions including the inferior parietal lobule and precuneus as well as the hippocampus and striatum. Perhaps the most consistent findings are recruitment of dorsal regions of the prefrontal cortex and posterior parietal regions, which together form the fronto-parietal network which is hypothesized to support manipulation of stimulus relations, information and executive functioning more broadly (Hinton et al., 2010; Ogawa et al., 2010; Schlund et al., 2007).

\*\*\*Insert Figure 1 About Here\*\*\*

Numerous clinical disorders are characterized by deficits in reasoning and the integration, ordering and relating of information (Avery, Williams, Woolard & Heckers, 2014; Onwuameze, Titone & Ho, 2016; Klabunde, Sagar, Hustyi, Kelley, Reiss & Hall, 2015; Solomon et al., 2015; Waltz et al., 2004). However, what is known is based largely on application of tasks such as TI or acquired equivalence paradigms (Schlichting & Preston, 2015). Expanding such paradigms to include AARR would provide a greater range of relations to explore and could generate new insights into dysfunctional relational abilities.

Regarding anxiety disorders, much of what is currently known about the neural circuits supporting negative affective responses and avoidance in humans and nonhumans is based on experiential learning which involves direct contact with aversive stimuli (Kirlic, Young, & Aupperle, 2017). Neuroscience research on AARR and symbolic generalization is uniquely positioned to examine the extent to which the same neural circuits are involved in supporting negative affective responses to, and avoidance of, derived relations. For example, findings showing that the amygdala responds to derived fear conditioned cues would represent an important extension of existing findings with learned fear cues and suggest an expanded functional role of the amygdala (see also, Paré and Quirk, 2017). Moreover, because AARR research is not easily performed with nonhumans, coupling AARR methods with brain imaging may provide insights into anxiety disorders not possible in nonhuman research.

In conclusion, we have sought to highlight some ways in which AARR applications in neuroscience and the broader clinical psychology and experimental psychopathology literatures might aid in further understanding the origin, acquisition, maintenance, and treatment of clinical disorders defined by excessive fear and avoidance. Historically, research on fear and avoidance has occurred within two separate research meta-theoretical traditions:

associative and operant (De Houwer, 2017). The two domains have not always interacted, although we firmly believe there is much to be gained for a contemporary translational understanding of fear, avoidance and anxiety disorders from research that combines both perspectives. Indeed, there is considerable potential for collaboration and cross-fertilization between cognitively oriented and more functionally oriented approaches to fear and avoidance. Here, we highlighted these areas of overlap and look forward to the next generation of functionally-informed, behavior analytic and experimental psychopathology research on AARR and fear and avoidance.



### References

- Acuna, B. D., Eliassen, J. C., Donoghue, J. P., & Sanes, J. N. (2002). Frontal and parietal lobe activation during transitive inference in humans. *Cerebral Cortex, 12*, 1312-1321.
- American Psychiatric Association (2013). *Diagnostic and Statistical Manual of Mental Disorders* (2<sup>nd</sup> ed.). Arlington, VA: American Psychiatric Publishing.
- Augustson, E. M., Dougher, M. J., & Markham, M. R. (2000). Emergence of conditional stimulus relations and transfer of respondent eliciting functions among compound stimuli. *The Psychological Record, 50*, 745-770.
- Avery, S. N., Williams, L. E., Woolard, A. A., & Heckers, S. (2014). Relational memory and hippocampal function in psychotic bipolar disorder. *European Archives of Psychiatry and Clinical Neuroscience, 264*, 199-211.
- Baldwin, D. S., Allgulander, C., Altamura, A. C., Angst, J., Bandelow, B., den Boer, J., Boyer, P., Davies, S., dell'Osso, B., Eriksson, E., Fineberg, N., Fredrikson, M., Herran, A., Maronm, E., Metspalu, A., Nutt, D., van der Wee, N., Vázquez-Barquero, J. L., & Zohar, J. (2010). Manifesto for a European anxiety disorders research network. *European Neuropsychopharmacology, 20*, 426-432.
- Barnes-Holmes, Y., Hussey, I., McEntegart, 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, and A. Biglan (Eds.), *The Wiley handbook of contextual behavioral science* (p.365-382). London: Wiley-Blackwell.
- Bennett, M., Hermans, D., Dymond, S., Vervoort, E., & Baeyens, F. (2015). From bad to worse: Symbolic equivalence and opposition in fear generalization. *Cognition & Emotion, 29*, 1137-1145. doi:10.1080/02699931.2014.97383

- Bennett, M., Meulders, A., Baeyens, F., & Vlaeyen, J. W. S. (2015). Words putting pain in motion: the generalization of pain-related fear within an artificial stimulus category. *Frontiers in Psychology, 6*:520. doi: 10.3389/fpsyg.2015.00520
- Bennett, M., Vervoort, E., Boddez, Y., Hermans, D., & Baeyens, F. (2015). Perceptual and conceptual similarities facilitate the generalization of instructed fear. *Journal of Behavior Therapy & Experimental Psychiatry, 48*, 149-155.
- Blackledge, J. T. (2007). Disrupting verbal processes: cognitive defusion in acceptance and commitment therapy and other mindfulness-based psychotherapies. *The Psychological Record, 57*, 555-576.
- Barrera, T. L., Szafranski, D. D., Ratcliff, C. G., Garnaat, S. L., & Norton, P. J. (2016). An experimental comparison of techniques: Cognitive defusion, cognitive restructuring, and in-vivo exposure for social anxiety. *Behavioural and Cognitive Psychotherapy, 44*, 249-254.
- Blackledge, J.T. (2015). *Cognitive defusion in practice: A clinician's guide to assessing, observing & supporting change in your client*. Oakland, CA: New Harbinger.
- Blackledge, J. T., Moran, D. J., & Ellis, A. E. (2009). Bridging the divide: Linking basic science to applied psychotherapeutic interventions - A relational frame theory account of cognitive disputation in rational emotive behavior therapy. *Journal of Rational-Emotive & Cognitive-Behavior Therapy, 27*, 232-248.
- Boddez, Y., Baeyens, F., Hermans, D., & Beckers, T. (2014). A learning theory approach to anxiety disorders: Human fear conditioning and the added value of complex acquisition procedures. In T. W. A. Ehring & P. Emmelkamp (Eds.), *International handbook of anxiety disorders: Theory, research and practice* (pp. 85-104). London: Wiley-Blackwell.

- Boddez, Y., Bennett, M., van Esch, S., & Beckers T. (2016). Bending rules: The shape of the perceptual generalization gradient is sensitive to inference rules. *Cognition & Emotion*. doi:10.1080/02699931.2016.1230541
- Boddez, Y., Davey, G., & Vervliet, B. (2016). Editorial: Experimental psychopathology: Defining the field. *Psychopathology Review*, 4, 109-111.
- Bouton, M. E., Mineka, S., & Barlow, D. H. (2001). A modern learning theory perspective on the etiology of panic disorder. *Psychological Review*, 108, 4-32. doi:10.1037/0033-295X.108.1.4
- Boyle, S., Roche, B., Dymond, S., & Hermans, D. (2016). Generalization of fear and avoidance along a semantic continuum. *Cognition & Emotion*, 30, 340-352. doi:10.1080/02699931.2014.1000831
- Clark, D. A., & Beck, A. T. (2010). *Cognitive therapy of anxiety disorders: Science and practice*. New York, NY: Guilford Press.
- Coelho, C. M., & Purkis, H. (2009). The origins of specific phobias: influential theories and current perspectives. *Review of General Psychology*, 13, 335-348. doi:10.1037/a0017759.
- Craske, M. G., Hermans, D., & Vansteenwegen, D. (2006). *Fear and learning: basic science to clinical application*. Washington, DC: APA Books.
- Craske, M. G., Rauch, S. L., Ursano, R., Prenoveau, J., Pine, D. S., & Zinbargh, R. E. (2009). What is an anxiety disorder? *Depression and Anxiety*, 26, 1066-1085. doi:10.1002/da.20633
- Craske, M.G., Treanor, M., Conway, C., Zbozinek, T., Vervliet, B. (2014). Maximizing exposure therapy: An inhibitory learning approach. *Behaviour Research and Therapy*, doi: 10.1016/j.brat.2014.04.006.

- Critchfield, T. S., & Reed, D. D. (2017). The fuzzy concept of applied behavior analysis research. *The Behavior Analyst*. doi:10.1007/s40614-017-0093-x
- De Houwer, J. (2017). A functional-cognitive framework for cooperation between functional and cognitive researchers in the context of stimulus relations research. *The Behavior Analyst*, doi:10.1007/s40614-017-0089-6
- De Houwer, J., Barnes-Holmes, D., & Barnes-Holmes, Y. (2017). What is cognition? A functional-cognitive perspective. In Steven C. Hayes and Stefan G. Hofmann (Eds.), *Core Processes of Cognitive Behavioral Therapies*. Oakland, CA: New Harbinger.
- Dickins, D. W., Singh, K. D., Roberts, N., Burns, P., Downes, J. J., Jimmieson, P., & Bentall, R. P. (2001). An fMRI study of stimulus equivalence. *Neuroreport*, *12*, 405-411.
- Dougher, M. J., Augustson, E., Markham, M. R., Greenway, D. E., & Wulfert, E. (1994). The transfer of respondent eliciting and extinction functions through stimulus equivalence classes. *Journal of the Experimental Analysis of Behavior*, *62*, 331-351. doi: 10.1901/jeab.1994.62-331
- Dougher, M. J., Hamilton, D. A., Fink, B. C., & Harrington, J. (2007). Transformation of the discriminative and eliciting functions of generalized relational stimuli. *Journal of the Experimental Analysis of Behavior*, *88*, 179–197. doi:10.1901/jeab.2007.45-05
- Dunsmoor, J. E., Kroes, M. C. W., Braren, S. H., & Phelps, E. A. (2017). Threat intensity widens fear generalization gradients. *Behavioral Neuroscience*, *131*, 168-175.
- Dunsmoor, J. E., & Murphy, G. L. (2015). Categories, concepts, and conditioning: how humans generalize fear. *Trends in Cognitive Sciences*, *19*, 73-77.
- Dunsmoor, J. E., Martin, A., & LaBar, K. S. (2012). Role of conceptual knowledge in learning and retention of conditioned fear. *Biological Psychology*, *89*, 300-305. doi:10.1016/j.biopsycho.2011.11.002

Dunsmoor, J. E., & Paz, R. (2015). Fear generalization and anxiety: behavioral and neural mechanisms. *Biological Psychiatry*, *78*, 336-343.

Dunsmoor, J. E., Prince, S. E., Murty, V. P., Kragel, P. A., & LaBar, K. S. (2011). Neurobehavioral mechanisms of human fear generalization. *NeuroImage*, *55*, 1878–1888. doi:10.1016/j.neuroimage.2011.01.041

Dymond, S., Molet, M., & Davies, L. (2017). The impact of arbitrarily applicable relational responding on evaluative learning about hypothetical money and shock outcomes. *Quarterly Journal of Experimental Psychology*, *70*, 1684-1699. doi: 10.1080/17470218.2016.1200639

Dymond, S., & Rehfeldt, R. (2000). Understanding complex behavior: The transformation of stimulus functions. *The Behavior Analyst*, *23*, 239-254.

Dymond, S., & Roche, B. (2009). A contemporary behavior analysis of anxiety and avoidance. *The Behavior Analyst*, *32*, 7-28.

Dymond, S., & Roche, B. (Eds.) (2013). *Advances in relational frame theory: Research & application*. Oakland, CA: New Harbinger.

Dymond, S., Dunsmoor, J. E., Vervliet, B., Roche, B., & Hermans, D. (2015). Fear generalization in humans: Systematic review and implications for anxiety disorder research. *Behavior Therapy*, *46*, 561-582. doi:10.1016/j.beth.2014.10.001

Dymond, S., Roche, B., & Bennett, M. (2013). Relational frame theory and experimental psychopathology. In Dymond, S., & Roche, B. (Eds.). *Advances in relational frame theory: Research & application* (pp.199-218). Oakland, CA: New Harbinger

Dymond, S., Roche, B., Forsyth, J. P., Whelan, R., & Rhoden, J. (2007). Transformation of avoidance response functions in accordance with the relational frames of same and opposite. *Journal of the Experimental Analysis of Behavior*, *88*, 249–262. doi:10.1901/jeab.2007.22-07

- Dymond, S., Roche, B., Forsyth, J. P., Whelan, R., & Rhoden, J. (2008). Derived avoidance learning: Transformation of avoidance response functions in accordance with the relational frames of same and opposite. *The Psychological Record, 58*, 271–288.
- Dymond, S., Schlund, M.W., Roche, B., De Houwer, J., & Freegard, G. (2012). Safe from harm: Learned, instructed, and symbolic generalization pathways of human threat-avoidance. *PLoS ONE, 7*(10): e47539. doi:10.1371/journal.pone.0047539
- Dymond, S., Schlund, M. W., Roche, B., & Whelan, R. (2014). The spread of fear: Symbolic generalization mediates graded threat-avoidance in specific phobia. *Quarterly Journal of Experimental Psychology, 67*, 247-259. doi:10.1080/17470218.2013.800124
- Dymond, S., Schlund, M., Roche, B., Whelan, R., Richards, J., & Davies, C. (2011). Inferred threat and safety: Symbolic generalization of human avoidance learning. *Behaviour Research and Therapy, 49*, 614-621. doi:10.1016/j.brat.2011.06.007
- Foa, E. B., & Kozak, M. J. (1986). Emotional processing of fear: Exposure to corrective information. *Psychological Bulletin, 99*, 20-35. doi:10.1037//0033-2909.99.1.20
- Foody, M., Barnes-Holmes, Y., Barnes-Holmes, D., & Luciano, C. (2013). An empirical investigation of hierarchical versus distinction relations in a self-based ACT exercise. *International Journal of Psychology and Psychological Therapy, 13*, 373-388.
- Gil-Luciano, B., Ruiz F. J., Valdivia-Salas, S., & Suarez-Falcon, J. C. (2016). Promoting psychological flexibility on tolerance tasks: Framing behavior through deictic/hierarchical relations and specifying augmental functions. *The Psychological Record, 66*, 1-9.
- Forsyth, J. P., & Eifert, G. H. (1996). The language of feeling and the feeling of anxiety: Contributions of the behaviorisms toward understanding the function-altering effects of language. *The Psychological Record, 46*, 607–649.

- 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. doi:10.1901/jaba.1998.31-137
- Gannon, S., Roche, B., Kanter, J. W., Forsyth, J. P., & Linehan, C. (2011). A derived relations analysis of approach-avoidance conflict: Implications for the behavioral analysis of human anxiety. *The Psychological Record, 61*, 227-252.
- Garcia-Guerrero, S., Dickins, T. E., & Dickins, D. W. (2014). The gradual extinction of transferred avoidance stimulus functions. *Psychological Record, 64*, 581-599. doi:10.1007/s40732-014-0062-7
- Grillon, C. (2008). Models and mechanisms of anxiety: evidence from startle studies. *Psychopharmacology, 199*, 421–437. doi:10.1007/s00213-007-1019-1.
- Guinther, P. M., & Dougher, M. J. (2015). The clinical relevance of stimulus equivalence and relational frame theory in influencing the behavior of verbally competent adults. *Current Opinion in Psychology, 2*, 21–25.
- Hayes, S. C. (1992). Verbal relations, time, and suicide. In S. C. Hayes and L. J. Hayes (Eds.), *Understanding verbal relations* (pp. 109-120). Reno, NV: Context Press.
- Hayes, S. C., & Hayes, L. J. (1992). Verbal relations, cognition, and the evolution of behavior analysis. *American Psychologist, 47*, 1383-1395. doi:10.1037/0003-066X.47.11.1383
- Hayes, S. C. (2004). Acceptance and Commitment Therapy, Relational Frame Theory, and the third wave of behavior therapy. *Behavior Therapy, 35*, 639-665.
- Hayes, S. C., Levin, M. E., Plumb-Villardaga, J., Villate, J. L., & Pistorello, J. (2013). Acceptance and commitment therapy and contextual behavioral science: Examining the progress of a distinctive model of behavioral and cognitive therapy. *Behavior Therapy, 44*, 180-198. doi:10.1016/j.beth.2009.08.002

- Hayes, S. C., Luoma, J. B., Bond, F. W., Masuda, A., & Lillis, J. (2006). Acceptance and commitment therapy: Model, processes and outcomes. *Behavior Research & Therapy*, *44*, 1-25.
- Hayes, S. C., Strosahl, K., & Wilson, K. G. (2011). *Acceptance and Commitment Therapy: The process and practice of mindful change* (2<sup>nd</sup> edition). New York: Guilford Press.
- Hermans, D., & Baeyens, F. (2013). Generalization as a basis for emotional change: Perceptual and non-perceptual processes. In D. Hermans, B. Rimé, & B. Mesquita (Eds.), *Changing emotions* (pp. 67-73). Hove, UK: Psychology Press.
- Hermans, D., Baeyens, F., & Vervliet, B. (2013). Generalization of acquired emotional responses. In M. D. Robinson, E. R. Watkins, & E. Harmon-Jones (Eds.), *Handbook of cognition and emotion* (pp. 117-134). New York: Guilford Press.
- Higgins, S. T., & Morris, E. K. (1984). Generality of free-operant avoidance conditioning to human behavior. *Psychological Bulletin*, *96*, 247-272.
- Hinton, E. C., Dymond, S., Von Hecker, U., & Evans, C. J. (2010). Neural correlates of relational reasoning and the symbolic distance effect: Involvement of parietal cortex. *Neuroscience*, *168*, 138-148.
- Hooper, N., & McHugh, L. (2013). Cognitive defusion versus thought distraction in the mitigation of learned helplessness. *The Psychological Record*, *63*, 209-217.
- Killeen, P. R., & Jacobs, K. W. (2016). Coal is not black, snow is not white, food is not a reinforcer: The roles of affordances and dispositions in the analysis of behavior. *The Behavior Analyst*, 1-22. doi:10.1007/s40614-016-0080-7
- Kirlic, N., Young, J., & Aupperle, R. L. (2017). Animal to human translational paradigms relevant for approach avoidance conflict decision making. *Behaviour Research and Therapy*, doi: 10.1016/j.brat.2017.04.010



- Klabunde, M., Sagar, M., Hustyi, K. M., Kelley, R. G., Reiss, A. L., & Hall, S. S. (2015). Examining the neural correlates of emergent equivalence relations in fragile X syndrome. *Psychiatry Research: Neuroimaging*, *233*, 373-379.
- Królicki, L., & Wróbel, A. (2011). A role for the right prefrontal and bilateral parietal cortex in four-term transitive reasoning: an fMRI study with abstract linear syllogism tasks. *Acta Neurobiologiae Experimentalis*, *71*, 479-495.
- Krypotos, A., Effting, M., Kindt, M., & Beckers, T. (2015). Avoidance learning: a review of theoretical models and recent developments. *Frontiers in Behavioral Neuroscience*, *9*, 189. doi:10.3389/fnbeh.2015.00189
- Lamal, P. A. (1998). Advancing backwards. *Journal of Applied Behavior Analysis*, *31*, 705-706.
- Lenaert, B., Boddez, Y., Griffith, J. W., Vervliet, B., Schruers, K., & Hermans, D. (2014). Aversive learning and generalization predict subclinical levels of anxiety: A six-month longitudinal study. *Journal of Anxiety Disorders*, *28*, 747-753.
- LeDoux, J. E. (2014). Coming to terms with fear. *Proceedings of the National Academy of Sciences of the United State of America*, *111*, 2871-2878. doi:10.1073/pnas.1400335111
- LeDoux, J. E., Moscarello, J., Sears, R., & Campese, V. (2016). The birth, death and resurrection of avoidance: A reconceptualization of a troubled paradigm. *Molecular Psychiatry*, *22*, 24-36. doi:10.1038/mp.2016.166
- Leslie, J. C., Tierney, K. J., Robinson, C. P., Keenan, M., Watt, A., & Barnes, D. (1993). Differences between clinically anxious and non-anxious participants in a stimulus equivalence training task involving threat words. *The Psychological Record*, *43*, 153-161.
- Levin, M. & Villatte, M. (2016). The role of experimental psychopathology and laboratory based intervention studies in contextual behavioral science. In S.C. Hayes, D. Barnes-

- Holmes, R. Zettle, & T. Biglan (Eds.), *Handbook of Contextual Behavioral Science* (pp. 347-364). Chichester, UK: Wiley Blackwell.
- Lewon, M., & Hayes, L. J. (2014). Toward an analysis of emotions as products of motivating operation. *Psychological Record, 64*, 813-825. doi:10.1007/s40732-014-0046-7
- Lissek, S. (2012). Toward an account of clinical anxiety predicated on basic, neurally mapped mechanisms of Pavlovian fear learning: The case for conditioned overgeneralization. *Depression and Anxiety, 29*, 257–263. doi:10.1002/da.21922
- Lissek, S., Bradford, D. E., Alvarez, R. P., Burton, P., Espensen-Sturges, T., Reynolds, R. C., & Grillon, C. (2014). Neural substrates of classically conditioned fear generalization in humans: A parametric fMRI study. *Social, Cognitive and Affective Neuroscience, 9*, 1134–1142. [http:// dx.doi.org/10.1093/scan/nst096](http://dx.doi.org/10.1093/scan/nst096)
- López-López, J. C. & Luciano, C. (2017). An experimental analysis of defusion interactions based on deictic and hierarchical framings and their impact on cognitive performance. *The Psychological Record*, doi:10.1007/s40732-017-0250-3
- Lonsdorf, T. B. et al. (2017). Don't fear 'fear conditioning': methodological considerations for the design and analysis of studies on human fear acquisition, extinction, and return of fear. *Neuroscience and Biobehavioral Review, 77*, 247–285.
- Luciano, C., Ruiz, F. J., Vizcaíno-Torres, R., Sánchez, V., Gutiérrez-Martínez, O., & López-López, J. C. (2011). A relational frame analysis of defusion interactions in Acceptance and Commitment Therapy. A preliminary and quasi-experimental study with at-risk adolescents. *International Journal of Psychology and Psychological Therapy, 11*, 165-182.
- Luciano, C., Valdivia-Salas, S., Ruiz, F. J., Rodríguez-Valverde, M., Barnes-Holmes, D., Dougher, M. J., Cabello, F., Sanchez, V., Barnes-Holmes, Y., & Gutierrez, G. (2013). Extinction of aversive eliciting functions as an analog of exposure to conditioned fear:

- Does it alter avoidance responding? *Journal of Contextual Behavioral Science*, 2, 120-134.
- Luciano, C., Valdivia-Salas, S., Ruiz, F. J., Rodríguez-Valverde, M., Barnes-Holmes, D., Dougher, M. J., López-López, J. C., Barnes-Holmes, Y., & Gutierrez-Martínez, O. (2014). Effects of an acceptance/defusion intervention on experimentally induced generalized avoidance: A laboratory demonstration. *Journal of the Experimental Analysis of Behavior*, 101, 94-111.
- Markham, R. G., & Markham, M. R. (2002). On the role of covarying functions in stimulus class formation and transfer of function. *Journal of the Experimental Analysis of Behavior*, 78, 509-524.
- Markham, M. R., Dougher, M. J., & Augustson, E. (2002). Transfer of operant discrimination and respondent elicitation via emergent relations of compound stimuli. *The Psychological Record*, 52, 325-350.
- Masuda, A., Hayes, S. C., Sackett, C. F., & Twohig, M. P. (2004). Cognitive defusion and self-relevant negative thoughts: Examining the impact of a ninety-year-old technique. *Behaviour Research and Therapy*, 42, 477-485. doi:10.1016/j.brat.2003.10.008
- Masuda, A., Hayes, S. C., Twohig, M. P., Drossel, C., Lillis, J., & Washio, Y. (2009). A parametric study of cognitive defusion and the believability and discomfort of negative self-relevant thoughts. *Behavior Modification*, 33, 250-262.
- May, R. J., Stewart, I., Baez, L., Freegard, G., & Dymond, S. (2017). Arbitrarily applicable spatial relational responding. *Journal of the Experimental Analysis of Behavior*, 107, 234-257. doi: 10.1002/jeab.250
- McEntegart, C., Barnes-Holmes, Y., Hussey, I., & Barnes-Holmes, D. (2015). The ties between a basic science of language and cognition and its clinical applications. *Current Opinion in Psychology*, 2, 56-59.

- Medina, M. A. L., Valverde, M. R., & Lopez, M. H. (2016). Transfer of conditioned fear-potentiated startle across equivalence classes: An exploratory study. *International Journal of Psychology and Psychological Therapy*, *16*, 249-263.
- Meulders, A., Vandebroek, N., Vervliet, B., & Vlaeyen, J. W. S. (2013). Generalization gradients in cued and contextual pain-related fear: An experimental study in healthy participants. *Frontiers in Human Neuroscience*, *7*, 345. doi: 10.3389/fnhum.2013.00345
- Meulders, A., & Vlaeyen, J. W. (2013). The acquisition and generalization of cued and contextual pain-related fear: An experimental study using a voluntary movement paradigm. *Pain*, *154*, 272-282. doi:10.1016/j.pain.2012.10.025
- Moors, A. (2014). Flavors of appraisal theories of emotion. *Emotion Review*, *6*, 303-307. doi:10.1177/1754073914534477
- Moss-Lourenco, P., & Fields, L. (2011). Nodal structure and stimulus relatedness in equivalence classes: Post-class formation preference tests. *Journal of the Experimental Analysis of Behavior*, *95*, 343-368.
- Ogawa, A., Yamazaki, Y., Ueno, K., Cheng, K., & Iriki, A. (2010). Neural correlates of species-typical illogical cognitive bias in human inference. *Journal of Cognitive Neuroscience*, *22*, 2120-2130.
- Onat, S., & Buchel, C. (2015). The neuronal basis of fear generalization in humans. *Nature Neuroscience*, *18*, 1811-1818. doi:10.1038/nn.4166
- Onwuameze, O. E., Titone, D., & Ho, B. C. (2016). Transitive inference deficits in unaffected biological relatives of schizophrenia patients. *Schizophrenia Research*, *175*, 64-71.
- Paré, D., & Quirk, G. J. (2017). When scientific paradigms lead to tunnel vision: lessons from the study of fear. *npj Science of Learning* *2*, doi:10.1038/s41539-017-0007-4
- Rachman, S. J. (1977). The conditioning theory of fear acquisition: A critical examination. *Behaviour Research and Therapy*, *15*, 375-387.

Ritzert, T. R., Forsyth, J. P., Berghoff, C. R., Barnes-Holmes, D., & Nicholson, E. (2015).

The impact of a cognitive defusion intervention on behavioral and psychological flexibility: An experimental evaluation in a spider fearful non-clinical sample. *Journal of Contextual Behavioral Science, 4*, 112-120.

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.

Rutherford, A. (2009). *Beyond the box: B.F. Skinner’s technology of behavior from laboratory to life, 1950s-1970s*. Toronto: University of Toronto Press.

Scheveneels, S., Boddez, Y., Vervliet, B., & Hermans, D. (2016). The validity of laboratory-based treatment research: Bridging the gap between fear extinction and exposure treatment. *Behaviour Research and Therapy, 86*, 87-94. doi: 10.1016/j.brat.2016.08.015.

Schlichting, M. L., & Preston, A. R. (2015). Memory integration: neural mechanisms and implications for behavior. *Current Opinion in Behavioral Sciences, 1*, 1–8.

Schlund, M. W., Hoehn-Saric, R., & Cataldo, M. F. (2007). New knowledge derived from learned knowledge: Functional-anatomical correlates of stimulus equivalence. *Journal of the Experimental Analysis of Behavior, 87*, 287-307.

Schlund, M. W., Cataldo, M. F., & Hoehn-Saric, R. (2008). Neural correlates of derived relational responding on tests of stimulus equivalence. *Behavioral and Brain Functions, 4*, 6.

Schlund, M. W., Brewer, A. T., Magee, S. K., Richman, D. M., Solomon, S., Ludlum, M., & Dymond, S. (2016). The tipping point: Value differences and parallel dorsal-ventral frontal circuits gating human approach-avoidance behavior. *NeuroImage, 136*, 94-105. doi:10.1016/j.neuroimage.2016.04.070

- Sidman, M. (1994). *Equivalence relations and behavior: A research story*. Boston: Authors Cooperative.
- Skinner, B. F. (1945). The operational analysis of psychological terms. *Psychological Review*, 52, 270–277.
- Skinner, B. F. (1953). *Science and human behavior*. New York: Free Press.
- Smyth, S., Barnes-Holmes, D., & Forsyth, J. P. (2006). Derived transfer of simple discrimination and self-reported arousal functions in spider fearful and non-spider-fearful participants. *Journal of the Experimental Analysis of Behavior*, 85, 223–246.
- Solomon, M., Ragland, J. D., Niendam, T. A., Lesh, T. A., Beck, J. S., Matter, J. C., Frank, M. J., & Carter, C. S. (2015). Atypical learning in autism spectrum disorders: a functional magnetic resonance imaging study of transitive inference. *Journal of the American Academy of Child & Adolescent Psychiatry*, 54, 947-955.
- Stewart, C., Stewart, I., & Hughes, S. (2016). A contextual behavioral approach to the study of (persecutory) delusions. *Journal of Contextual Behavioral Science*, 5, 235–246.
- Taylor, I., & O'Reilly, M. F. (1997). Toward a functional analysis of private verbal self-regulation. *Journal of Applied Behavior Analysis*, 30, 43–58. doi:10.1901/jaba.1997.30-43.
- Tyndall, I., Papworth, R., Roche, B., & Bennett, M. (2017). Differential effects of word-repetition rate on cognitive defusion of believability and discomfort of negative self-referential thoughts postintervention and at one-month follow-up. *The Psychological Record*, doi:10.1007/s40732-017-0227-2
- Rodriguez-Valverde, M. R., Luciano, C., & Barnes-Holmes, D. (2009). Transfer of aversive respondent elicitation in accordance with equivalence relations. *Journal of the Experimental Analysis of Behavior*, 92, 85–111 doi: 10.1901/jeab.2009.92-85.

- Vansteenwegen, D., Vervliet, B., Iberico, C., Baeyens, F., Van den Bergh, O., & Hermans, D (2005). The repeated confrontation with videotapes of spiders in multiple contexts attenuates renewal of fear in spider-anxious students. *Behaviour Research & Therapy*, *45*, 169-79.
- Vervliet, B., Vansteenwegen, D., & Eelen, P. (2004). Generalization of extinguished skin conductance responding in human fear conditioning. *Learning & Memory*, *11*, 555–558. doi:10.1101/lm.77404
- Vervliet, B., & Raes, F. (2013). Criteria of validity in experimental psychopathology: Application to models of anxiety and depression. *Psychological Medicine*, *43*, 2241-2244. doi:10.1017/S0033291712002267
- Vervoort, E., Vervliet, B., Bennett, M., & Baeyens, F. (2014). Generalization of human fear acquisition and extinction within a novel arbitrary stimulus category. *PLoS ONE*, *9*, doi:10.1371/journal.pone.0096569.
- Vilardaga, R., Hayes, S. C., Levin, M. E., & Muto, T. (2009). Creating a strategy for progress: A contextual behavioral science approach. *The Behavior Analyst*, *32*, 105-133.
- Villatte, M. Villatte, J. L., & Hayes, S. C. (in press). A reticulated and progressive strategy for developing clinical applications of RFT. *The Psychological Record*.
- Villatte, J. L., Vilardaga, R., Villatte, M., Vilardaga, J. C. P., Atkins, D. C., & Hayes, S. C. (2016). Acceptance and Commitment Therapy modules: Differential impact on treatment processes and outcomes. *Behaviour Research and Therapy*, *77*, 52-61.
- Waltz, J. A., Knowlton, B. J., Holyoak, K. J., Boone, K. B., Back-Madruga, C., McPherson, S., Masterman, D., Chow, T., Cummings, J. L., & Miller, B. L. (2004). Relational integration and executive function in Alzheimer's disease. *Neuropsychology*, *18*, 296-305.

Wittchen, H.-U., Jönsson, B., & Olesen, J., (2005). Editorial: towards a better understanding of the size and burden and cost of brain disorders in Europe. *European*

*Neuropsychopharmacology*, *15*, 355–356.

Zeithamova, D., Dominick, A. L., & Preston, A. R. (2012). Hippocampal and ventral medial prefrontal activation during retrieval-mediated learning supports novel inference.

*Neuron*, *75*, 168-179.



**Figure Caption**

*Figure 1.* The distributed brain network supporting various forms of derived responding in humans (see text for details). PFC = prefrontal cortex; DLPFC = dorsolateral PFC; VLPFC = ventrolateral PFC. (Whelan, R., & Schlund, M. W. (2013). Reframing relational frame theory research: Gaining a new perspective through the application of novel behavioral and neurophysiological methods. In S. Dymond and B. Roche (Eds.), (2013) *Advances in relational frame theory: research and application*, pp. 151-177. Reprinted with permission from New Harbinger.)