12-14-2016

Laboratory Analogue Investigation of Cognitive Defusion and Cognitive Reappraisal Strategies In the Context Of Symbolically Generalized Avoidance

Matthew Donati

LABORATORY ANALOGUE INVESTIGATION OF COGNITIVE DEFUSION AND COGNITIVE REAPPRAISAL STRATEGIES IN THE CONTEXT OF SYMBOLICALLY GENERALIZED AVOIDANCE

by

MATTHEW DONATI

Under the Direction of Akihiko Masuda, PhD

ABSTRACT

The present study used a basic behavioral paradigm derived from Relational Frame Theory (RFT), a contemporary behavioral account of language and cognition, to validate principle-based definitions of the cognitive interventions of defusion and reappraisal. Ninety-one participants first underwent an RFT learning paradigm that established symbolically generalized avoidance. Participants were then randomized to a defusion, reappraisal, or control condition. The main outcomes were equivalence responding—indicative of the trained relational network and analogous to the content of cognition—and avoidance—analogous to the behavioral impact of cognition. Defusion and reappraisal significantly reduced avoidance responding, providing support for the hypothesis that these interventions target the behavior of relational framing. Mediation analyses, conducted on an exploratory basis, revealed differences between the mechanisms of defusion and reappraisal and provided preliminary support for the classification of these interventions as a functional context intervention and a relational context intervention, respectively.

INDEX WORDS: Defusion, Reappraisal, Relational Frame Theory, Acceptance and Commitment Therapy, Cognitive Behavioral Therapy, Fusion

LABORATORY ANALOGUE INVESTIGATION OF COGNITIVE DEFUSION AND COGNITIVE REAPPRAISAL STRATEGIES IN THE CONTEXT OF SYMBOLICALLY GENERALIZED AVOIDANCE

by

MATTHEW DONATI

A Thesis Submitted in Partial Fulfillment of the Requirements for the Degree of

Master of Arts

in the College of Arts and Sciences

Georgia State University

2016

LABORATORY ANALOGUE INVESTIGATION OF COGNITIVE DEFUSION AND COGNITIVE REAPPRAISAL STRATEGIES IN THE CONTEXT OF SYMBOLICALLY GENERALIZED AVOIDANCE

by

MATTHEW DONATI

Committee Chair: Akihiko Masuda

Committee: Dominic Parrott

Erin Tone

Electronic Version Approved:

Office of Graduate Studies

College of Arts and Sciences

Georgia State University

December 2016

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

1 INTRODUCTION

1.1 Overview

Acceptance and Commitment Therapy (ACT; Hayes, Strosahl, & Wilson, 2012) is a contextual cognitive behavioral therapy (Hayes, Villatte, Levin, & Hildebrandt, 2011) which has been gaining increasing support as an empirically validated treatment. One of the distinguishing features of ACT is its approach to problematic cognitions. ACT cognitive defusion strategies aim to alter the behavioral *functions* of target cognitive events, such as negative self-referential thoughts. For example, in a widely used defusion strategy, rapid verbal repetition, individuals repeat a target thought out loud until the words comprising the thought lose their psychological impact.

ACT proponents consider defusion strategies to be distinct from more traditional cognitive strategies, such as cognitive restructuring or cognitive reappraisal. However, the extant evidence cannot be used to determine if these strategies operate through theoretically distinct or consistent mechanisms. The debate continues mainly because these applied concepts/techniques have not yet been empirically examined at a principles level. The present reverse-translational analogue experiment examined and compared cognitive defusion and cognitive reappraisal at the level of the basic behavioral principles of verbal behavior (i.e. language and cognition), as conceptualized by Relational Frame Theory (RFT, Hayes, Barnes-Holmes, & Roche, 2001), a contemporary behavioral account of verbal behavior.

In this rationale, I will first briefly introduce the key assumptions of Contextual Behavioral Science (CBS; Hayes, Barnes-Holmes, & Wilson, 2012), the research program responsible for both RFT and ACT, in the context of the behavioral analysis of verbal behavior. I will discuss the defining principles of RFT, and how RFT delineates the contextual control of verbal behavior into the *relational context* and the *functional context*. According to RFT, the relational context controls the specifics of verbal behavior, such as thought content, while the functional context controls the impact of verbal behavior, such as avoidance. Following the discussion of RFT, I will introduce the applied constructs of cognitive defusion and cognitive reappraisal, and a preliminary RFT conceptualization of these strategies. More specifically, I will argue that both strategies may work similarly at the principles level by altering the *functional context* of verbal behavior. Finally, I will present the reverse-translational analogue experiment, which was designed to test this hypothesis by using a widely studied RFT laboratory behavioral avoidance paradigm.

1.2 Contextual Behavioral Science

Behavior analysis and verbal behavior. Throughout the history of behavior analysis, verbal behavior has presented a unique challenge (Hayes, 2004). A preliminary definition of ***verbal behavior,*** otherwise referred to as ***symbolic behavior***, is behavior that involves language and/or cognition (Hayes, Barnes-Holmes, & Roche, 2001). Traditional behavior analysts, including B.F. Skinner (Skinner, 1957), attempted to model verbal behavior in terms of direct contingencies (i.e., operant and respondent principles). A ***direct contingency*** account defines behavior strictly in terms of functional relations that have been established by contingent contact in an individual's learning history, such as when a discriminative stimulus (e.g. a blue light) consistently signals the availability of a subsequent reinforcer (e.g. a food pellet will release if a bar is pressed). A behavior's ***function*** refers to the pattern of its interrelations with contextual variables. Such variables can make the behavior more likely to occur in the future (reinforcers), can signal when it is advantageous to perform the behavior (discriminative stimulus), or increase motivation to perform the behavior (establishing operations), among other things. These interrelations among behaviors and their contexts constitute ***functional relationships.***

Direct contingency accounts also include generalization based on formal (i.e. perceptual) stimulus properties; for instance, a child frightened by a fluffy white dog may subsequently respond with wariness in the presence of other fluffy white animals or objects. ***Indirect contingency*** accounts, in contrast to direct contingency accounts, also allow for the possibility that stimuli can come to be related according to symbolic dimensions without having been in contingent contact in an individual's learning history. For instance, if the child in the last example is told that a new neighborhood "contains many fluffy white dogs," she may fear that neighborhood even though she has never witnessed fluffy white dogs in that neighborhood.

The direct contingency accounts of verbal behavior failed to yield well-specified, applied techniques, which were previously the hallmark of behavior analysis (Hayes & Wilson, 1993). This is in part because they could not account for functional relationships that had no basis in an individual's direct learning history. One example of a relationship that direct contingency accounts provide poor explanations for is when an individual develops a phobia, despite having no prior traumatic experiences with the feared stimuli (Dymond & Roche, 2009; Lovibond, 2006).

Cognitive Behavioral Therapy model. With the failure of the behavior analysts to provide an adequate account of verbal behavior, theorists began to abandon the strictly inductive approach of behavior analysis and to allow for cognitive mediating constructs in their theories (Hayes, Villatte, Levin, & Hildebrandt, 2011). The traditional Cognitive Behavioral Therapies (CBTs) arose largely from this effort. Traditional CBT models assume that problematic cognitions cause emotional and behavioral actions (Herbert & Forman, 2009; Hofmann, Asmundson, & Beck, 2013). Therefore, the cognitive techniques of CBT, including cognitive restructuring and cognitive reappraisal, focus on the form and frequency of thinking. These techniques typically attempt to alter problematic thinking patterns (e.g. making global negative thoughts more circumscribed) or reduce the frequency of occurrence of negative thinking patterns.

Contextual Behavioral Science. Contextual Behavioral Science (CBS; Hayes, Barnes-Holmes, et al., 2012) offers an alternate framework with which to understand verbal behavior. CBS arose out of a contemporary attempt to analyze and clinically treat verbal behavior utilizing the contextual strategies of behavior analysis (Biglan & Hayes, 1997; Hayes, Barnes-Holmes, et al., 2012; Zettle, 2011). Although CBS, like the CBT model, considers cognition an important factor in the etiology of psychopathology, it differs from the CBT model in its underlying assumptions (e.g. what it takes for a factor to be considered causal), which stem from a distinct philosophy of science. CBS researchers espouse a scientific philosophy of functional contextualism. Proponents of mechanistic or descriptive philosophies of science assume one reality that they attempt to describe by mapping their theories and concepts with observations. They determine which scientific principles are "true" according to how well the proposed principles correspond to the one reality. In contrast, ***functional contextualism*** is a pragmatic approach which eschews ontological concerns (i.e. is there one reality or does reality depend on a given purpose?; can reality ever be directly measured?) and chooses scientific principles according to how well the proposed principles contribute to a stated goal (Biglan & Hayes, 1997).

A piece of knowledge, a proposed behavioral principle for example, is considered to be "true" to the extent that it aides in *prediction* and *influence* of the phenomenon of interest, with *precision*, *scope* and *depth*. Precise principles can predict and influence individual events using a limited number of concepts. Principles with scope apply to a wide range of topographically distinct events. Principles with depth are in accord with useful constructs at other levels of analysis (Hayes & Berens, 2004). Since any manipulation to achieve prediction and influence must occur in the behavioral context, analyses define behavior in terms of functional relations with manipulable contextual variables, and behaviors and variables are defined in terms of their function rather than their topography.

Unlike traditional behavioral analysis, in which applied techniques emerged largely from basic experimental findings (Hayes & Berens, 2004), CBS espouses a reticulated development strategy (Hayes, Barnes-Holmes, et al., 2012), in which research at different levels of analysis (e.g. basic, applied) progresses in parallel. This strategy prevents research at one level from stalling while waiting on relevant work from the other. While each level progresses according to its own standards, the priority placed on *depth* dictates that scientists working at each level must strive to connect their research finding at other levels over the long term (Hayes, Barnes-Holmes, et al., 2012).

CBS approach to verbal behavior. As noted earlier, CBS researchers share the assumption of many CBT models that cognition and behavior are causally related (S C Hayes, Barnes-Holmes, et al., 2012; S C Hayes, 2004; Zettle & Hayes, 1986). However, the act of thinking is viewed as a behavior which requires a functional analysis to understand, like any outward behavior. Therefore, CBS researchers view the relationship between thoughts and outward behavior as a behavior-behavior chain. As such, the analysis of outward behavior that is influenced by cognition is not complete until empirical work has identified both the contingencies that give rise to thinking and, perhaps even more important, the contextual factors that link the act of thinking to acts of outward behavior (Zettle & Hayes, 1986; Zettle, 2011). As mentioned above, within CBS, RFT is a contemporary behavioral theory that defines language and cognition in terms of functional relationships with manipulable contextual variables at the basic level (Hayes, Barnes-Holmes, et al., 2001). At the applied level, Acceptance and Commitment Therapy (ACT; Hayes, Strosahl, & Wilson, 2012) is a RFT-informed treatment approach that attempts to identify and alter the contextual factors that give rise to the causal relationship between thoughts and behavior (Hayes et al., 1999).

1.3 A CBS basic model of verbal behavior

Sidman (1971) published a seminal study demonstrating that an individual learning disabilities was able to indirectly learn relations, between printed words and pictures and between pictures and spoken words, that were never directly trained. Since this work appeared in the literature, laboratory experiments have demonstrated that when humans who can use language are taught patterns of discrimination among stimuli, they will demonstrate patterns of responding to those stimuli that are not accounted for by the direct training, but are nevertheless coherent and predictable. This behavior, derived from the directly trained patterns, is hypothesized to be at the heart of human symbolic behavior (Hayes, 1994).

Consider first that if a parent wanted to teach a child that the word “ball” stands for and means an actual physical ball, the parent might use a conditional discrimination procedure, and reinforce the child for choosing a ball rather than another object when presented with the word “ball”. As illustrated just above, even very young children and many animals can learn to conditionally discriminate objects in this way; however, mastering this discriminative process does not mean that they have learned how to read the word “ball” or even grasped the word’s meaning. For example, when presented with a ball, a young child or an animal who routinely selects the correct object in response to the word "ball" would not choose the word "ball" to label the object at a greater probability than chance. That is, the trained discrimination performance would not reverse, a necessary property of language use (Sidman & Tailby, 1982).

Relational responding. Symbolic behavior involves more than conditional discriminations between comparison and sample stimuli (Sidman & Tailby, 1982). RFT posits that symbolic behavior requires the fundamental ability to respond according to the relations among stimuli (Hayes, Blackledge, & Barnes-Holmes, 2001). RFT defines ***relational responding*** as responding to a stimulus according to its relation with another stimulus (or stimuli), as opposed to a direct learning history with the former. Both young humans and animals demonstrate the simplest version of this ability. Rhesus monkeys, for example, can be trained to always choose the larger of two sticks (Harmon, Strong, & Pasnak, 1982). Such training involves multiple trials in which a larger stick is introduced each time and the subject is reinforced for choosing the larger stick. Once a monkey learns the behavior of choosing the larger stick, that monkey will choose a larger stick that it had no prior experience with, over a smaller stick that it had been reinforced for choosing in previous trials. In other words, the subject's behavior comes under the control of the relation ("bigger than"), rather than the history of reinforcement for choosing any one stick. After being presented with numerous trials using multiple exemplars, individuals will learn to abstract the aspect of their reinforced performance that is constant: the relation between the reinforced and non-reinforced choices.

Derived relational responding. According to RFT, language-able humans take this behavior a step further. When they learn to respond according to a relation between stimuli, they *derive* relations that were not directly reinforced. Firstly, if a human subject learns that A *is more than* B, then she will derive that B *is less than* A. RFT refers to this bi-directionality of human relational responding as ***mutual entailment***. Secondly, if a human being is taught that A *is more than* B and B *is more than* C, then she will derive that A *is more than* C, and C *is less than* A. RFT refers to this combined bi-directionality and transitivity of human relational responding as ***combinatorial mutual entailment***.

The applicability of these performances to language can best be seen with the equivalence relation. In the case of the equivalence relation, the derived performances described above would fulfill the requirements, from mathematical theory, for one stimulus to symbolize another (Sidman & Tailby, 1982). These properties are identity (A = A), reflexivity (if A = B, then B = A), and transitivity (if A = B, and B = C, then A = C). In addition to equivalence, RFT experimenters have demonstrated that individuals are able to respond according to a range of relations, such as better than, part of, and greater than (Barnes-Holmes, Hayes, Dymond, & O'Hora, 2001). These basic relations provide the necessary components for the analysis of such complex behaviors as problem solving and analogical thinking (Hayes, Gifford, Townsend, & Barnes-Holmes, 2001; Stewart, Barnes-Holmes, Hayes, & Lipkens, 2001).

Arbitrary applicability. As described above, animals and non-language-able humans can respond to relations established according to physical properties of relata (e.g. the size of sticks). However, in order to effectively use language and symbolism, one must be able to learn and apply relations that are not defined in terms of physical dimensions. A person needs this ability, for example, in order to understand that letters on a page or utterances of sounds stand for worldly objects. Consider the example provided by Hayes, Luoma, Bond, Masuda, and Lillis, (2006) of a verbally competent child who has experience using a nickel, but not a dime and who is told that a dime *is more than* a nickel. The phrase *is more than* will function as a relational cue that will result in the dime having functions of value for the child and a relative preference for dimes over nickels. This occurs despite the fact that a nickel is physically larger than a dime, because the relation is ***arbitrarily applicable***. That is, the relational response does not depend on the physical properties of the objects or ideas being linked and can instead be cued by socially determined aspects of the context (in this example, the spoken words "is more than").

Transformation of stimulus function. According to RFT, all behavioral functions, including those that are psychological (e.g., fear, bias, avoidance), can come under the control of relational responding (Hayes, Fox, et al., 2001). For example, Dougher, Hamilton, Fink, and Harrington (2007) demonstrated that arbitrary stimuli (i.e. shapes and nonsense words) can come to elicit aversion for subjects as a result of relational responding. The researchers experimentally established the following relations (A > B, A < C), and demonstrated that participants derived the following combinatorial mutually-entailed relations (B < C, C > B) (for the purpose of clarity, alpha-numeric labels will be used to represent experimental stimuli throughout this proposal). They then paired B with electric shock. Over time, B acquired a new function: it began to elicit aversion, as measured by skin conductance. C was then presented over repeated trials. Participants demonstrated greater skin conductance in response to C than B, even though C had never been paired with electric shock and had never been directly paired with B. Thus, the experimenters demonstrated that the behavioral functions of B transformed the stimulus function of C in accordance with the derived "greater" than relation. RFT refers to this aspect of relational responding as a ***transformation of stimulus function***.

Relational and functional contexts. The type of relations and the transformation of function according to these relations can be controlled by separate aspects of the behavioral context (Wulfert & Hayes, 1988). The ***relational context*** specifies the type of relation between stimuli (e.g. coordination – the stimuli are equated; comparison – the stimuli differ from each other in a specific way, etc.) and is often abbreviated as $\mathbf{C_{rel}}$ in the RFT literature. The ***functional context*** determines what functions are transformed according to the relation and is referred to as $\mathbf{C_{func}}$. Most stimuli will have numerous possible functions, and the functions that will be in effect in any one instance of behavior depend on the C_{func}. For example, some version of the following exercise is often used to illustrate the dynamic process of relational context and functional context in human linguistic practice.

> If you have never learned the word "sitruuna," then reading the word will have very limited functions for your behavior, mostly related to its physical properties like the arrangement of the lines and shapes on the page. However, if you are told that a sitruuna is a lemon (in fact it is the Finnish word for lemon), then a relation of coordination will be established between sitruuna and lemon. Now consider the different effects of the two following statements: 1) Imagine the color of a bright ripe sitruuna; 2) Imagine biting into a big juicy sitruuna. The first statement may lead to the mental image of a yellow lemon, while the latter may lead to puckering or mouth-watering. In an RFT analysis, the difference in psychological functions of the word sitruuna in these two statements is attributed to the surrounding words that establish the C_{func} and specify the transformation of function. While the transformation of function has been altered, according to RFT, the C_{rel} including the specified relata has not been altered. In each of the statements given, sitruuna is in the same relation of coordination with a lemon.

Relational framing. The preceding sections describe characteristics of verbal behavior as observed in laboratory experiments and categorized by RFT researchers. RFT researchers propose that responding to a stimulus based on its relations to other stimuli, as opposed to a direct learning history with the stimulus, is the functional unit underlying all verbal acts (i.e. all acts of verbal behavior can be functionally defined in terms of relations and the contingencies controlling relating). The behavior of relating does not require additional behavioral principles to explain. Rather, RFT researchers view verbal activity as operant behavior – behavior controlled by surrounding contingencies (e.g., antecedants that signal the availability of reinforcers, the reinforcers themselves, etc). The key novel contributions of RFT are perhaps recognizing that relating can be considered a class of operant behavior and functionally defining the characteristics of this class. A class of operant behavior (i.e. high-order class, generalized class) is a group of behaviors that are controlled more by the contingencies arranged for the shared characteristics of the group, as opposed to the contingencies for one particular instance of the behavior. For example, the behavior of a child imitating the seatbelt use of an adult for the first time is the result of past reinforcement for other imitation behaviors.

RFT researchers propose that verbal behavior can best be defined as an operant class of a particular way of relating stimuli. The essential characteristics of this relating are arbitrary applicability (i.e. applicability regardless of physical properties) mutual entailment (i.e. bi-directionality of relations), combinatorial mutual entailment (i.e. tranisitivity of relations), and the transformation of stimulus function. The operant class of behavior that exhibits these characteristics is known as ***arbitrarily applicable derived relational responding***, and is otherwise referred to as ***relational framing***. The term ***relational frame*** is used as shorthand to summarize the analysis of any one instance of relational framing, but what is being referred to is a behavioral process rather than a fixed construct. In addition to the particular stimuli involved, a relational frame is specified by the type of relation (C_{rel}; e.g. comparison, coordination, etc.) and the function transformed (C_{func}; e.g. avoidance elicitation, mental imagery, etc.).

There is evidence to suggest that relational framing is indeed an overarching operant class of behavior learned in early childhood, in the same way that imitation is learned as an overarching operant by way of training with multiple exemplars (Berens & Hayes, 2007). Relational framing abilities coincide with language skills in children and learning disabled adults, and training in relational framing improves language abilities (Hayes, Levin, & Plumb-Vilardaga, 2013).

RFT analysis of an applied problem: verbally established avoidance. RFT paradigms have been developed to study numerous behaviors, including discrimination, sexual arousal, and approach/avoidance (Dymond, May, Munnelly, & Hoon, 2010)**.** Out of these, the most widely studied is the approach/avoidance paradigm (Dymond et al., 2010), likely because avoidance is a central concept in in the CBS theoretical conceptualization of psychopathology (Chawla & Ostafin, 2007). The RFT model of verbally established approach/avoidance provides a behavioral account of the overgeneralization of avoidance observed in populations with a range of anxiety disorders. People in these populations often fear and avoid neutral stimuli despite the absence of an aversive conditioning history (Ahmed & Lovibond, 2014). CBS adherents hold that this overgeneralization occurs according to symbolic dimensions of similarity between stimuli, and thus, sometimes call it ***symbolically generalized avoidance*** (Dymond, Schlund, Roche, De Houwer, & Freegard, 2012).

RFT readily explains how stimuli that have never been directly paired with adverse consequences can come to obtain avoidance-eliciting functions (Augustson & Dougher, 1997). Dymond et al. (2011) experimentally demonstrated this process. First, they established two equivalence classes, each with three arbitrary stimuli ([A1, B1, C1] and [A2, B2, C2]). An ***equivalence class*** is a group of arbitrary stimuli which are established as functionally equivalent by relational frames of coordination among the members of the class. Equivalence classes are commonly established by using match to sample (MTS) procedures, in which a sample stimuli is provided and participants are reinforced for choosing one of several comparison stimuli. In the case of Dymond et al. (2011), a *one-to-many* training procedure was used in which the following performances were reinforced (for each of the two equivalence classes): A -> B, A -> C ("A->B", reads, "given A, choose B"). The MTS training functions as a contextual cue (C_{rel}) for individuals to relationally frame the stimuli in the class as equivalent. To demonstrate that this has occurred, combinatorial mutual entailment is confirmed by testing for the following MTS performances in the absence of reinforcement: B -> C, C -> B.

After establishing the equivalence classes, Dymond et al. (2011) differentially trained avoidance and approach functions for the B1 and B2 stimuli using emotionally laden images and sounds. They established B1 as a CS^+ and B2 as a CS^-. Finally, responding to the C1 and C2 stimuli was tested in extinction. Participants avoided the C1 stimuli, while approaching the C2 stimuli. This demonstrated that the directly trained functions of the two B stimuli had transformed the functions of the two C stimuli in accordance with the established relational frames. Participants avoided the C1 stimuli despite the fact that it had never been paired with an aversive outcome or directly related to B1. The differential transformation of function to the two C stimuli can be attributed to relational framing. Several other studies have succeeded in demonstrating the same process (Dymond, Roche, Forsyth, Whelan, & Rhoden, 2008; Dymond, Schlund, Roche, & Whelan, 2014; Luciano et al., 2014; Roche, Kanter, Brown, Dymond, & Fogarty, 2008).

Treatment implication. The analysis of verbal behavior in terms of relational framing has significant treatment implications (Törneke, 2010; Wilson, Hayes, Gregg, & Zettle, 2001)

When extending RFT research to clinical situations, the analogy of spoken verbal behavior is applied to cognition. Thoughts are viewed as private verbal behavior comparable to self-talk, in which the speaker and the listener are the same (Törneke, 2010). Within this framework, the C_{rel} controls the form and frequency of thinking, and the C_{func} controls the behavioral impact of thinking. Given that the C_{rel} of verbal behavior is manipulable separate from the C_{func}, RFT predicts that it is possible to alter the impact of thinking without necessarily changing the content of thoughts (e.g., Zettle & Hayes, 1986). This is arguably the most significant clinical implication of RFT.

1.4 Applied research on problematic verbal behavior

At the applied level, ACT researchers proposed the concept of ***cognitive fusion*** to capture the dimension of the behavioral impact of thinking as distinct from the content of thinking. This conceptual account is partially informed by RFT (Hayes, Strosahl, & Wilson, 1999), but has also developed from behavioral analytic accounts of client-therapist interactions (Zettle & Hayes, 1986; Zettle, 2011). In ACT literature, the concept of ***cognitive fusion*** refers to a *context* wherein a person's private events (primarily thoughts) are taken literally and readily responded to as if they are in themselves the real world referents, as opposed to verbal events (Hayes, Strosahl, Wilson, 2012). This may occur when a person's attention is focused on the content of thoughts without awareness of the process of thinking itself and "buys into" the supposed truth or meaning of thoughts (Hayes, 2004). Researchers have attempted to delineate specific aspects of the cognitive fusion context (Blackledge, 2007; Hayes, Barnes-Holmes, et al., 2001) and have developed manipulations with the expressed goal of targeting this context.

Features of the cognitive fusion context. ACT researchers posit that the social-verbal community determines much of the fusion context. For example, the ***context of reason giving***

refers to the reinforcement by the social community of behavior that implies cognitions are acceptable causes, or reasons, for outward behavior (Zettle & Hayes, 1986). When individuals are reinforced for giving private events as reasons for their behavior, the likelihood of those private events leading to behavioral outcomes in the future increases (Zettle & Hayes, 1986).

Other elements of the fusion context include the ubiquitous, socially-determined features of communication required for the ready use of language. These are believed to include an attentional focus on language products, the stimulus products of relational framing, as opposed to the process of language-ing itself. Seamless communication also requires consistent use of accepted language parameters, such as consistent rates and frequencies of words spoken and syntactic structures (Blackledge, 2007). Together these aspects form the ***context of literality*** and result in indirect stimulus functions of words and thoughts (presumably the functions resulting from relational framing) to transfer readily and dominate over other behavioral contingencies (Hayes, Strosahl, & Wilson, 1999).

Cognitive defusion exercises. ***Cognitive defusion*** is the opposite of cognitive fusion and refers to a context in which thoughts do not elicit automatic behavioral reactions, but instead, are directly experienced as psychological events that may or may not elicit a behavioral response (Blackledge, 2007; Hayes et al., 2006). When considering cognitive defusion strategies, it is necessary to explicitly outline the goal of clinical treatment (Blackledge, 2007; Wilson & Murrell, 2004). Unlike many therapies, ACT does not espouse the goal of symptom reduction. As a functional contextual treatment, the goal of ACT is successful working (Hayes, 2004). At the level of analysis of treating individuals, successful working is defined as an increase in movement towards (or a decrease in movement away from) the individual's chosen values, with values being those qualities of living that an individual finds most important to engender in daily life (Hayes, Strosahl, et al., 2012; Wilson & Murrell, 2004). Cognitive defusion interventions can therefore be defined as strategies to alter the behavioral impact of thinking, regardless of the form and frequency of thinking, for the purpose of increasing valued living (Hayes, Strosahl, et al., 2012; Wilson & Murrell, 2004).

Many cognitive defusion techniques are designed to accomplish defusion by altering the socio-verbal context of therapy sessions (Hayes et al., 2006; Hayes, Strosahl, et al., 2012; Masuda, Hayes, Sackett, & Twohig, 2004). Perspective taking exercises in ACT, for example, are designed to redirect a client's focus from the content of thoughts to the process of thinking itself, in order to create psychological distance from thoughts (Blackledge, 2007). In one such exercise, thoughts are written on note cards and treated as tangible objects, and clients explore the difference between holding thoughts closely, blocking their view of valued outcomes, versus holding thoughts at a distance. "Description-evaluation" exercises (Hayes & Smith, 2005) demonstrate the discrepancies between language and real world referents. In the "teach me how to walk" exercise (Hayes & Smith, 2005), clients notice that the words describing the process of walking cannot possibly capture the complexity and depth of the experience (Blackledge, 2007).

Still other ACT defusion exercises are designed to violate the language parameters of the context of literality. One such exercise, rapid vocal repetition, is the cognitive defusion technique that has been studied the most extensively (Masuda et al., 2010, 2004; Masuda, Hayes, & Twohig, 2009). In this exercise, participants are asked to think about what comes to mind when they hear the word milk. This likely includes the psychological functions of imagining milk, such as a cold temperature, white color and creamy texture. Participants are then instructed to repeat the word rapidly at a high frequency over an approximately 30 second time period. When the word is repeated in this way, language parameters are violated and indirect functions are disrupted in favor of direct functions. The psychological functions of the word milk fall away and what is experienced is the direct experience of the word's component sounds (Blackledge, 2007).

Defusion evidence. ACT outcome studies, have demonstrated the efficacy of a treatment which draws heavily from the theoretical principles of defusion and which employs defusion techniques (Öst, 2014; Swain, Hancock, Hainsworth, & Bowman, 2013). Yet, this evidence provides only limited validation of the theoretical constructs. In order to demonstrate correspondence between theory, treatment techniques, proposed mechanisms, and outcomes, we must turn to laboratory based studies and treatment mediation and moderation studies (Herbert, Gaudiano, & Forman, 2013; Levin, Hildebrandt, Lillis, & Hayes, 2012). Several laboratory analogue studies have demonstrated the efficacy of isolated defusion interventions with nonclinical populations. Defusion interventions for the treatment of negative thoughts have been significantly more effective at reducing distress associated with the target thought than control and thought distraction conditions (Healy et al., 2008; Hooper & McHugh, 2013; Masuda et al., 2010, 2004; Pilecki & Mckay, 2012) and as effective as cognitive restructuring conditions (Deacon, Fawzy, Lickel, & Wolitzky-Taylor, 2011; Yovel, Mor, & Shakarov, 2014). In addition to demonstrating that defusion techniques reduce the distress surrounding negative cognitions, the majority of these studies have attempted to provide evidence that defusion works through theoretically consistent processes. To do so they have operationalized the function of thoughts, independent from the form or frequency of thoughts, with self-reported *metacognitive beliefs* about thoughts. Defusion interventions have reduced self-reported believability of thoughts, accuracy of thoughts and perceived importance of thoughts (Deacon et al., 2011; Masuda et al., 2010, 2004; Yovel et al., 2014). Furthermore, mediation studies have provided self-report evidence that changes in cognitive defusion mediate ACT treatments for depression (Zettle, Rains, & Hayes, 2011), psychosis (Bach, Gaudiano, Hayes, & Herbert, 2012), tinnitus distress (Hesser, Westin, Hayes, & Andersson, 2009), anxiety (Arch, Wolitzky-Taylor, Eifert, & Craske, 2012) and mixed anxiety and depression (Forman, Chapman, et al., 2012).

Limitations of the defusion construct. Together, the laboratory and mediation studies suggest that defusion is a useful construct and target for intervention. However, beyond that, the conclusions that can be drawn from the existing research are limited. The limitations of the defusion research include the methodologies employed and variability in both definition and operationalization of the construct.

The evidence in support of the defusion construct is mostly self-report in nature (Masuda et al., 2010, 2004). This is the case for many constructs in psychology, but it becomes particularly problematic for the construct of defusion. As of yet, psychometrically sound and widely accepted scales for measuring defusion are still lacking, and the majority of defusion studies used visual analogue scales (Deacon, Fawzy, Lickel, & Wolitzky-Taylor, 2011). The lack of commonly accepted measures likely has to do with the subtlety of the defusion construct. When responding, individuals are required to differentiate between some event occurring, such as a thought or a distressing feeling, and the impact of that event, such as avoidance or metacognitive beliefs about the event. It may be difficult for participants to make this distinction considering that individuals do not readily differentiate between distressing experiences and the behavioral impact of that distress on self-report measures (Gamez, Kotov, & Watson, 2010). Consequently, a major concern of recently developed defusion measures is that they confound understanding of defusion with endorsement of defusion (Forman, Herbert, et al., 2012; Gillanders et al., 2014).

It becomes particularly difficult to determine the theoretical relevance of the aforementioned self-report evidence when it is considered with regard to the goal of treatment. As mentioned above, in order to demonstrate successful working, defusion strategies need to contribute to valued living. Laboratory based studies are unlikely to capture the clinical outcomes of valued living. Nevertheless, there are design options that may better proximate clinical phenomenon relevant to valued living, such as experimental designs that allow for the direct observation of behavioral patterns that are theorized to impede valued living in clinical populations. One such behavioral pattern is symbolically generalized avoidance (defined above). In the CBS model of psychopathology, it is theorized that the pervasive avoidance observed in clinical populations is often the result of the symbolic generalization of avoidance contingencies (Dymond & Roche, 2009; Dymond et al., 2011, 2014). Symbolically generalized avoidance functions to limit valued living when it overrides competing approach contingencies, at which point it is considered maladaptive or problematic from a CBS perspective (Forsyth, Eifert, & Barrios, 2006). The in-laboratory observation of symbolically generalized avoidance with competing approach contingencies would be a better approximation of values-limiting behavior and provide a more relevant test that defusion works as hypothesized.

Further qualifying the evidence for defusion, there is a large amount of variability in the way researchers have attempted to both define and operationalize the construct. In addition to the definitions presented above, defusion has alternatively been described as creation of psychological distance (Forman, Herbert, et al., 2012), perspective taking regarding thoughts (Gillanders et al., 2014), and acceptance of thoughts (Luciano et al., 2014). To operationalize defusion, Zettle and Hayes (1986) compared cognitive attributions to validity of attributions; Masuda et al., (2010, 2004) and Healy et al. (2008) used believability as a proxy for defusion;

while Deacon et al. (2011) used both purported accuracy of thoughts and importance of thoughts to operationalize defusion.

The ways in which defusion has been defined, operationalized, and measured make it difficult to determine whether defusion is working through mechanisms consistent with CBS theory and whether these mechanisms are distinct from related constructs, specifically those of cognitive restructuring and cognitive reappraisal. As the research currently stands, there appears to be considerable overlap between defusion and cognitive reappraisal in particular (Mennin, Ellard, Fresco, & Gross, 2013; Yovel et al., 2014). Mennin and colleauges (2013) suggest that much of the evidence purportedly in support of the theory of defusion (e.g., the function of thoughts can be changed separately from their content) could as easily be interpreted in terms the emotion processing model underlying cognitive reappraisal.

Cognitive reappraisal. Cognitive reappraisal is a cognitive change (i.e. changing the content of thoughts) strategy that involves intentionally reinterpreting emotion-eliciting contexts in order to alter habitual emotional and behavioral reactions (Gross, 1998). Cognitive reappraisal constitutes a common procedure utilized across various CBT protocols for the treatment of emotional disorders (Wilamowska et al., 2010).

The rationale for cognitive reappraisal techniques derives from reappraisal theories of emotion (Troy, Wilhelm, Shallcross, Amanda, & Mauss, 2010) which posit that it is an individual's evaluation of a context, including ascribing verbal meaning to the situation, that determines an emotional response, rather than the situation itself (Lazarus & 1984; Ortony, Clore, & Collins, 1988; Scherer, 1988). Gross' emotion processing model (1998, 2002) is a contemporary attempt to capture the consensus among the reappraisal theories of emotion. In this model, cognitive evaluations of stimuli (e.g. emotion cues) mediate contact with the stimuli and the subsequent emotional response. Evaluations initiate coordinated physiological, affective, and behavioral response tendencies that are adaptively appropriate for the perceived situation, and different evaluations initiate distinct response tendencies. Once activated, these response tendencies can be modulated, and the modulated response is akin to the fully evinced emotional and behavioral reaction. This model highlights two time points where behavioral strategies can be introduced to regulate emotion. Antecedent-focused strategies intervene at the point of the cognitive evaluation, prior to the full initiation of the emotional, behavioral and physiological responses; and consequence-focused strategies intervene at the point of the response modulation, after full activation of the response tendencies (Gross, 2002).

Cognitive reappraisal is theorized to be an antecedent-focused emotion regulation strategy (Gross, 1998). When faced with aversive emotion-eliciting contexts, individuals are encouraged to adopt a different perspective on what they observe and experience. They may be asked to view a situation more objectively or to view it as a passive observer would. The situation then takes on verbal meaning different (e.g. less threatening) from the habitual interpretation and activates a different (often less reactive) response tendency. For example, before showing subjects a film chosen to elicit disgust (an arm amputation), Gross (1998) instructed subjects to “try to *think* about what you are seeing objectively, in terms of the technical aspects of the events you observe…try to *think* about what you are seeing in such a way that you don't feel anything at all. [emphasis added]” (p.227). One way participants may respond to these instructions is to think about the film clip as a medical professional would, noting the technical aspects of the procedure. In this way the same stimuli may take on very different functions for the person’s behavior, perhaps eliciting curiosity and increased attending, as opposed to disgust and avoidance.

Several analogue studies have been performed to study the effect of cognitive reappraisal techniques. Gross (1998) and Richards & Gross (2000) demonstrated that reappraisal decreased the stress response and emotional arousal of participants when viewing distressing film clips. Hofmann, Heering, Sawyer, & Asnaani (2009) found that reappraisal reduced anxiety related to giving an impromptu speech, significantly more than a control condition and equal to an acceptance condition. Limited evidence also suggests that cognitive reappraisal may mediate associations between treatment and successful outcomes. Goldin et al. (2012) demonstrated that increases in self–efficacy in the domain of cognitive reappraisal mediated CBT treatment outcomes for individuals with social anxiety disorder.

Comparisons between reappraisal and defusion. Although defusion is often presented by CBS proponents as a set of strategies wholly distinct from CBT strategies, several researchers have noted that cognitive defusion and reappraisal may work according to the same underlying principles (Arch & Craske, 2008; Herbert & Forman, 2013; Mennin et al., 2013). Cognitive defusion and cognitive reappraisal both attempt to change the stimulus function of contacted stimuli (Yovel et al., 2014). In addition, they both may rely on changes in cognitive distancing (Arch & Craske, 2008; Mennin et al., 2013; Yovel et al., 2014), perspective taking (Mennin et al., 2013; Yovel et al., 2014), or both to do so.

Thus far, no studies have directly compared defusion and reappraisal interventions, but two studies have compared defusion to cognitive restructuring, a technique closely related to reappraisal. These studies did not find differences between defusion and restructuring for the treatment of negative cognitions, in terms of efficacy or process as measured by self-report (Deacon et al., 2011; Yovel et al., 2014). However, as has been the case for most studies comparing ACT and CBT processes, the studies in question were designed with a focus on head-to-head comparative efficacy, and such designs are not well suited for the testing of theory or the validation of constructs (Kanter, 2013; Levin et al., 2012).

The current limitations of the applied defusion research—difficult to interpret self-report evidence and overlap with related constructs—appear to be representative of the limits of applied research without strong links to basic principles. The existing research base can be furthered by empirically validating defusion at the level of underlying principles (Blackledge, 2007). This can be accomplished by utilizing the theoretically specified paradigms of RFT to examine defusion techniques in a laboratory setting (Blackledge, Moran, & Ellis, 2008). Applying these paradigms to related techniques, such as cognitive reappraisal, would also help validate defusion by better establishing the boundaries of the construct (Clark & Watson, 1995). Furthermore, this effort would help to increase interdisciplinary communication between ACT and CBT researchers (Mennin et al., 2013).

1.5 Theoretical integration of applied and basic research

C_{func} vs. C_{rel} manipulations. In an initial attempt to link the RFT and ACT lines of research, researchers have offered conceptualizations of cognitive treatments in terms of RFT principles. As detailed above, when conceptualizing cognition in terms of RFT principles, thinking is viewed as an act of relational framing. The content of thoughts consists of arbitrary stimuli which are framed with real world referents (or other verbal stimuli) and gain their behavioral impact via the transformation of stimulus functions. From a contextual behavioral standpoint, the options for intervention all involve altering some portion of the context. Therefore, interventions are viewed as targeting the relational context (C_{rel}), the functional context (C_{func}), or both.

Strictly speaking, most cognitive interventions, including defusion and reappraisal, are verbally mediated, and therefore occur in the relational context (Hayes et al., 2013; Hayes, Strosahl, et al., 2012). Still, interventions occurring in the relational context can achieve changes in C_{func}, and they can be expected to differ in their overall effect on the C_{rel}, C_{func}, or both, of existing relations, depending on the focus of these interventions. According to RFT, cognitive treatments are categorized as *focusing on* or *targeting* either the C_{rel} or the C_{func} of thinking, because the distinction is hypothesized to be of clinical significance (Hayes, 2004; Hayes et al., 2006, 1999).

RFT predicts that targeting the C_{func} of existing relational networks is more effective than targeting the C_{rel}, because the properties of relational framing, including the properties of extinction, dictate that existing relations are difficult to alter, once formed (Wilson & Hayes, 1996). In addition, attempts to do so can have untoward effects (Hayes et al., 2013; Hayes et al., 2006; Hayes, Strosahl, et al., 2012). Studies of extinction reveal that successfully extinguished responses will return to performance, depending on the post-extinction context (Bouton, 2004). Therefore, extinction does not involve "unlearning," but rather, new learning which inhibits the extinguished responses (Bouton, 2002). Viewing symbolic relating as operant behavior suggests that relations are never erased. Indeed, Wilson and Hayes (1996) demonstrated that punished relations (e.g. a given relational network of events) do not disappear, but reemerge in new contexts. Moreover, the arbitrary applicability of relational framing vastly broadens the potential contexts that can elicit a given relational response, making it unlikely that relational responses can be effectively extinguished (Hayes et al., 2013). Finally, the contingencies hypothesized to reinforce relational framing, such as deriving coherent relational networks (i.e. making sense) are so pervasive as to make relational framing practically self-reinforcing (Alisha M Wray, Dougher, Hamilton, & Guinther, 2012). Together these properties suggest that existing thought patterns cannot be erased or eliminated and are virtually impossible to effectively control.

Interventions focused on altering existing relational networks could have paradoxical effects. For one, trying to directly change the content of thinking supports the idea that thoughts are in themselves responsible for emotions and behavior (Hayes, 2004), which could increase the behavioral regulatory function of thinking (S C Hayes et al., 2006; Zettle & Hayes, 1986). Said another way, C_{rel} interventions can become ineffective if the focus on thought content adversely impacts $C_{func.}$ This can occur, for example, when a focus on the importance of thought content leads to thought suppression (Hooper, Saunders, & McHugh, 2010), considered a detrimental outcome by numerous theoretical orientations (Gross, 2002; Hofmann et al., 2009; Wegner, 1994; Wolgast, Martin, Lundh, Lars-Gunnar, Viborg, 2012).

Defusion and reappraisal as C_{func} manipulations. Cognitive defusion interventions are focused on altering the functional impact of thinking, regardless of the content or frequency of thinking, and thus are conceptualized as C_{func} interventions (Hayes et al., 2013; Hayes et al., 2006; Hayes, 2004). Specifically, it is hypothesized that cognitive defusion strategies work by disrupting, fully or partially, the transformation of stimulus functions between thoughts and their real world referents (Blackledge, 2007). In contrast, cognitive treatment techniques, at least as originally conceptualized, focus on directly changing thoughts and beliefs, in form and frequency (Hofmann et al., 2013). As such, cognitive therapy strategies are generally conceptualized as C_{rel} manipulations focused on eliminating or altering existing relational networks (Hayes et al., 2013; Hayes et al., 2006; Hayes, 2004).

However, the focus on thought content of cognitive therapy strategies does not preclude these strategies from achieving beneficial functional context manipulations. The analogue evidence in favor of cognitive restructuring and cognitive reappraisal provides support for this possibility (Deacon et al., 2011; Gross, 1998; Gross, 2002; Yovel et al., 2014). Cognitive restructuring, for example, could achieve beneficial C_{func} outcomes by encouraging clients to view thoughts objectively as outward objects might be viewed (Arch & Craske, 2008; Mennin et al., 2013). Cognitive reappraisal, which does not necessarily rely on challenging thoughts, may be a particularly effective C_{func} manipulation (Hayes, Strosahl, et al., 2012). Studies reveal that even though reappraisal focuses on thought content, it does not lead to thought suppression and reduces the emotional impact of stimuli, a beneficial C_{func} outcome (Arch & Craske, 2008; Gross, 1998; Gross, 2002).

It is too early to speculate on how reappraisal works at the level of specific relational frames, but one possibility is that reappraisal works by increasing cognitive flexibility in response to emotion-laden stimuli (Hayes, Strosahl, et al., 2012). Cognitive flexibility refers to the general cognitive ability to inhibit narrow, non-workable habitual responding in favor of holding multiple ideas about a target stimuli (Johnco, Wuthrich, & Rapee, 2014). Allowing for multiple, often discrepant, relations may bring awareness to thoughts as verbal stimuli (Hayes, Strosahl, et al., 2012). In RFT terms, when there are discrepant or competing stimulus functions indirectly related to the same stimulus, responding according to any one of these indirect stimulus functions may be reduced in favor of responding according to direct stimulus functions of the stimulus (Gavin & Roche, 2008).

1.6 Conceptualization of cognitive strategies as C_{func} manipulations

The self-report evidence provides some support for the idea that cognitive defusion techniques create a de-synchrony between relational products (i.e. thoughts occurring) and their behavioral effects (Hayes et al., 2004; Masuda et al., 2010; Zettle, et al., 2011). However, without utilizing controlled laboratory procedures to establish relational framing, it is not possible to determine whether interventions are manipulating the verbal transformation of function (Blackledge, 2007). One research study published to date has attempted to utilize an RFT laboratory paradigm to empirically validate the conceptualization of cognitive defusion as a C_{func} manipulation (Luciano et al. 2014). No studies to date have utilized an RFT paradigm to study other cognitive strategies.

Luciano et al. (2014) utilized a between groups design to compare the effect of a defusion/acceptance/motivation protocol (DEF) on the derived transformation of avoidance functions, to that of a motivation only protocol (MOT) and a post hoc control group (CMOT). They employed the same basic procedures used by Dymond et al. (2011) to first establish two equivalence classes consisting of arbitrary stimuli and to then train relationally derived avoidance functions for one member of one class, the C1 stimuli. Participants were then exposed to one of three treatments. The MOT treatment lasted approximately 5 minutes, during which the experimenter read a script that notified participants that non avoidance in the presence of arbitrary stimuli would be rewarded with points possibly worth a small dollar value. The DEF treatment consisted of the MOT protocol plus didactic and experiential defusion, acceptance, and perspective taking exercises conducted by the experimenter and lasted approximately 15-20 minutes. The CMOT treatment consisted of the MOT exercises plus task irrelevant exercises to control for the amount of experimenter interaction and lasted approximately 15-20 minutes. After exposure to one of the three protocols, participants were presented with two tasks. The first task tested for avoidance of the C1 stimuli. The second tested for the equivalence relations trained in the first phase of the experiment. The percentage of participants who did not avoid were 100%, 40%, and 20% in the DEF, CMOT, and MOT conditions, respectively. All participants demonstrated relational responding in line with the equivalence classes trained. This experiment demonstrated that the DEF treatment altered the verbal transformation of stimulus function of stable relational networks. Said another way, the intervention successfully altered the C_{func} of experimentally established verbal relating without altering the existing relational networks.

Luciano et al. (2014) has limitations. The primary limitation of this study is that the treatment manipulation lacks construct validity. The DEF protocol was not well specified and employed a number of ACT strategies. This limits the relevance of these results to the analysis of defusion techniques. In addition, the study employed a post-hoc control group. Combined with the small sample sizes (10 per group) and differential attrition between the post hoc control group and the other groups this potentially limits the internal validity of the results. Lastly, Luciano et. al. (2014) did not compare the defusion condition to related cognitive strategies.

In the present study, I extended the work of Luciano et al. (2014) by employing a highly valid defusion manipulation, rapid vocal word repetition (Masuda et al., 2010, 2004, 2009). In addition, I utilized the same RFT procedures to test the conceptualization of cognitive reappraisal as a C_{func} manipulation. The approach/avoidance paradigm was chosen because this paradigm has been utilized the most frequently in RFT research (Dymond & Roche, 2009; Dymond et al., 2011, 2008, 2012, 2014; Luciano et al., 2013; Roche et al., 2008) and verbally established avoidance is a behavior of primary clinical importance in the CBS research agenda (Chawla & Ostafin, 2007).

1.7 Summary

The RFT model of symbolic behavior postulates that the process of cognitive interventions can be understood in terms of manipulations of the functional context (C_{func}) and the relational context (C_{rel}) of relational framing. Principles of operant behavior suggest that C_{func} manipulations are preferable to C_{rel} manipulations. Designed as C_{func} manipulations, cognitive defusion techniques are efficacious for altering metacognitive beliefs regarding problematic thoughts, as measured by self-report. However, there is a gap between the RFT work and the applied defusion literature. Specifically, defusion lacks a well specified and validated principles-based definition. As a result, there is considerable overlap between defusion and other related constructs, such as cognitive reappraisal; and the applied evidence cannot be readily utilized to extend theory.

1.8 Hypotheses

In the present study, there were three intervention groups: cognitive defusion, cognitive reappraisal, and a control. As described below in detail, the primary outcomes were avoidance behavior in the presence of neutral stimuli, which had been relationally coordinated with aversive stimuli, and outcome equivalence responding (i.e. match to sample responses). Outcome avoidance responses demonstrated relational framing with the correspondent transformation of avoidance eliciting functions, and equivalence responding exhibited in tact relational networks. A C_{func} manipulation is expected to reduce avoidance responses by disrupting or otherwise altering the transformation of function without the need to disrupt existing relational networks. Therefore, I expected to see reduced avoidance responses without changes in equivalence responding in both the defusion and reappraisal conditions, as compared to the control condition.

Since defusion techniques are specifically designed as C_{func} manipulations, I expected that the defusion technique would be relatively more efficacious than the reappraisal intervention. A commonly studied defusion exercise, rapid vocal word repetition, was used in order to maximize the construct validity of the defusion intervention. The reappraisal exercise was modelled after that utilized by Gross (1998, 2002). To the extent that these exercises validly operationalize defusion and reappraisal, respectively, I can draw conclusions about the defining principles of these constructs.

Research aim 1: Do cognitive defusion and cognitive reappraisal alter the stimulus functions of verbal stimuli (i.e., an avoidance-evoking function) established by relational framing? Is defusion relatively more efficacious in this regard?

> **Hypothesis 1**: Group assignment (defusion, reappraisal, or control) will have a significant effect on avoidance behavior, in the presence of stimuli with indirectly trained aversive functions, with a significantly greater reduction in the defusion and reappraisal condition as compared to the control condition and in the defusion condition as compared to the reappraisal condition.

Research aim 2: In the treatment of the symbolic generalization of avoidance do defusion and reappraisal function as C_{func} interventions? Is defusion relatively more efficacious in this regard?

> **Hypothesis 2**: Group membership will moderate the relationship between outcome avoidance, a measure of the transformation of stimulus functions, and outcome equivalence resonding, a measure of the relational network. Said another way, defusion and reappraisal will result in a relative de-synchrony between equivalence responding and avoidance. The moderation effect will be stronger in the defusion condition as compared to the reappraisal condition.

2 METHOD

2.1 Participants

Ninety-eight participants were recruited from undergraduate psychology courses at Georgia State University for partial fulfillment of mandatory research credits through SONA, an online research participant pool. They were undergraduate students at least 18 years of age who could read and write English. Written informed consent was obtained from all subjects.

All 98 participants completed questionnaires, including a demographics questionnaire and the BAS/BIS scales. Of these 98 participants, 7 were excluded from the laboratory procedures of the study due to a programming error. The errant code muted the volume of the program, invalidating the paradigm, depending on the results of a random number generator. It randomly affected approximately 1 out of every 6 participants. The error was corrected between the 43rd and 44th participant. All affected participants were dismissed from the laboratory procedures before reaching randomization into groups. These 7 participants were excluded from the attrition rate calculations, to allow for comparison to other studies.

Of the remaining 91 participants, 60 completed all of the experimental procedures of the study (34% attrition; see below for a detailed breakdown of attrition). These 60 participants constituted the final sample.

Participants in the final sample were generally traditional undergraduate students, aged 18 to 26 years (M = 19.18, SD = 1.82). The sample was predominantly female (n = 53; 88.3%) and was racially diverse. When permitted to identify with more than one racial/ethnic category, 58.3% identified as African American/Black; 23.3% identified as White; 18.3% identified as Asian or Pacific Islander, and 13.3% identified as Latino/Hispanic.

The group of participants who did not complete the study (non-completers) was similar demographically to the final sample. The sample of non-completers was predominantly female (*n* = 29, 93.5%). When permitted to identify with more than one racial/ethnic category, 58.1% identified as African American/Black; 16.1% identified as White; 22.6% identified as Asian or Pacific Islander, 9.7% identified as Latino/Hispanic, and 3.2% identified as Native American/American Indian. The average age of non-completers (*M* = 22.00, *SD* = 9.47) differed somewhat from that of completers due to the ages of three non-traditional students in the non-completers group. When these participants were excluded from the mean age calculation, the average age of non-completers (*M* = 19.13) was very similar to that of completers.

2.2 Materials and design

Apparatus and setting. Participants were seated at a table in a small experimental room (4m x 4m) containing a Dell™ OptiPlex 980 Windows PC with a 14-inch display. A computer program written in Java 8™ controlled all stimulus presentations and recorded all participant responses. All participant responses were made using either the computer's external keyboard or mouse.

Design overview. Figure 2.1 depicts the general experimental procedure. Pre-intervention training and testing procedures established equivalence responding and symbolically generalized avoidance responding in the presence of arbitrary stimuli. Participants were then randomly assigned to receive either one of two interventions or a control procedure. In the post intervention phases, participants were again tested for equivalence responding and avoidance responding.

The pre-intervention training and testing phases largely followed the same general procedures of more recent RFT studies that have trained relational framing behavior in the laboratory (Dymond et al., 2011, 2012; Luciano et al., 2014). In order to demonstrate arbitrarily applicable avoidance responding these studies followed three main phases: 1) formation of equivalence classes (most often 2 three-member classes); 2) establishing avoidance in the presence of one stimuli from one class (the B1 stimuli) using operant conditioning or a combination of respondent and operant conditioning; and 3) demonstration of the transformation of avoidance responding to a related stimuli (the C1 stimuli). Studies differ in the training method used to establish equivalence classes (e.g., match to sample, relational completion procedure), the type of arbitrary stimuli (e.g., shapes, random combinations of letters), and the unconditioned stimuli used to establish avoidance responding (e.g. aversive photos and sounds, electric shock). Studies also vary widely in the training and testing criterion used.

The present study followed the same general pre-intervention phases of equivalence class formation, avoidance conditioning, and transformation of avoidance. The majority of the methodology was modelled after the procedures utilized by Dymond and colleagues (2011, 2012) and Luciano and colleagues (2013, 2014). Training and testing criterion were arrived at using the recommendations from these studies and pilot data.

Similar to Luciano et al. (2014), the post intervention outcome phases are largely a repeat of the pre-intervention testing phases (equivalence and avoidance responding), with the addition of a response cost for avoidance.

One of the main points of divergence in the procedures of the present study is in the choice of interventions (as discussed in detail above). In addition, the choice of a time delay response cost in the post intervention portion of the study is unique in this type of experiment. In comparison, Luciano et al. (2014) used loss of points as a response cost. The use of made up words as arbitrary stimuli also differed from many RFT studies, which mainly utilized strings of random letters and geometric shapes. Lastly, the presentation of the arbitrary stimuli using audio output in addition to visual output (made possible by the use of made up words) differs from previously used procedures.

Arbitrary stimuli: Although the majority of RFT studies use abstract shapes or random three letter combinations as arbitrary stimuli, it was felt that the psychological realism of the cognitive interventions could be increased if the stimuli resembled real words. Hence, the arbitrary stimuli consisted of non-words from Massaro, Venezky, & Taylor (1979), which have been utilized in a small number of previous RFT experiments (Hooper et al., 2010; Whelan, Cullinan, Donovan, & Valverde, 2005). The words meet several English language criteria: (a) being orthographically regular; (b) being pronounceable; (c) containing common vowel and consonant spellings; and (d) having no more than three letters for a medial consonant cluster if one occurs (e.g. "boceem", "remond"; Hooper et al., 2010; Massaro et al., 1979). For the sake of clarity, stimuli are labeled herein using alphanumerics (A1, B1, C1, A2, B2, C2) that signify their experimentally trained functions—their positions in the equivalence classes and their trained/derived functions. The stimuli were randomized across the alphanumeric positions for each participant. Participants never saw the alphanumeric labels. All stimuli were presented simultaneously in written format on the computer screen and in audio format over headphones.

On the computer screen stimuli were presented in black font on a white background in 24 point lowercase Arial font. Over headphones stimuli were presented at a 1 decibel volume level. The digital audio recordings used by the computer program were made using a neutral accent and the pronunciations from Massaro et al. (1979).

Aversive and non-aversive stimuli. Visual and auditory stimuli were selected from the International Affective Picture System (IAPS; Lang, Bradley, & Cuthbert, 2008) and the

International Affective Digitized Sounds (IADS; Bradley, Lang, Margaret, & Peter, 2007) databases for use as aversive and non-aversive unconditioned stimuli during the avoidance conditioning and testing phases. A total of 20 photographs, 10 aversive (e.g., bodily mutilations) and 10 non aversive (e.g., landscapes), as well as 10 aversive sounds (e.g., a female screaming) were selected based on their use as unconditioned stimuli in past RFT experiments (Dymond et al., 2011, 2008). The auditory stimuli were presented via headphones at a 1 decibel volume level. The visual stimuli were presented on the computer monitor in a size of 600 X 800 pixels.

Interventions. Participants were randomly assigned using block randomization to receive one of two interventions—defusion , reappraisal—or a control exercise. The defusion intervention (see Appendix A) consisted of a rapid vocal repetition exercise and was modelled after that used by Masuda and colleagues (2010, 2004, 2009). The defusion rationale and technique was presented to participants by an experimenter. There was a practice phase in which the participant performed rapid vocal repetition using the word" milk" and an intervention phase in which the participant performed the exercise using the C1 stimulus. (The C1 stimulus is the neutral stimulus to which avoidance functions were transferred via the symbolic relation with the B1 stimulus. Avoidance of the C1 stimulus is the main experimental outcome.) The intervention took approximately 7.5 minutes to complete.

The reappraisal intervention (see Appendix B) was modelled after the procedures used by Gross (1998) and Keng, Robins, Smoski, Dagenbach, & Leary (2013). The reappraisal rationale was presented by the experimenter followed by a practice phase in which participants were asked to imagine an emotionally-valenced situation and use the reappraisal strategy of reacting to the situation in an objective manner with a detached and unemotional attitude. Following the practice, participants were asked to repeat the C1 stimulus once out loud while employing the same reappraisal strategy. Similar to the defusion condition, the intervention took approximately 7.5 minutes to complete.

The control condition consisted of a task-irrelevant exercise (see Appendix C) lasting approximately 7.5 minutes similar to that used by Luciano et al. (2014) which controlled for time and the amount of experimenter interaction.

2.2 Measures

Demographics questionnaire. Participants filled out a questionnaire with information on gender orientation, race/ethnicity identities, and age.

Dispositional behavioral inhibition and behavioral activation. Individual differences in behavioral inhibition sensitivity and behavioral activation sensitivity were measured using Carver's and White's (1994) BIS/BAS scales. The scales are based on Gray's Reinforcement Sensitivity Theory, which hypothesizes the neurobiological systems of the Behavioral Inhibition System (BIS) and the Behavioral Activation System (BAS; Gray, 1970). The BAS is responsible for activating responses in the presence of reward or escape contingencies, while the BIS is responsible for inhibitory responses when aversive consequences are signaled (Gray, Hanna, Gillen, & Rushe, 2016). The BIS/BAS consists of 20 items across four subscales using a four-point Likert-like rating response (ranging from "1 = very true for me", to "4 = very false for me"). Three scales measure different dimensions of the BAS: BAS-Drive measures motivation to pursue goals (e.g., "When I want something I usually go all-out to get it."); BAS-Fun Seeking measures sensitivity to the possibility of novel rewards (e.g., "I will often do things for no other reason than that they might be fun"); and BAS-Reward Responsiveness measures reactivity to rewarding consequences ("When good things happen to me, it affects me strongly"). The BIS subscale measures sensitivity to aversive consequences (e.g., "I feel pretty worried or upset when

I think or know somebody is angry at me"). BIS/BAS scores have been related to the occurrence of anxiety, depression, and externalizing disorders (Johnson, Turner, & Iwata, 2003). Of particular relevance to the present study, baseline levels of the BIS/BAS scales are predictive of college student responses to approach-avoidance conflicts (Berkman, Lieberman, & Gable, 2009).

In the present study, baseline levels of BAS and BIS sensitivity, as measured by the BIS/BAS scales, were considered as potential covariates when modeling participant avoidance responding. Considering study completers only, Cronbach's alpha for BAS-Drive ($\alpha = .74$) and BIS ($\alpha = .70$) were within an acceptable range, while Cronbach's alpha values for BAS-Fun Seeking ($\alpha = .65$) and BAS-Reward Responsiveness ($\alpha = .69$) fell below acceptable levels, according to common research standards. The low Cronbach's alpha of the two scales may have been due to the low number of items in each scale, as Cronbach's alpha is sensitive to the number of scale items (Cortina, 1993). However, results from these scales in the present sample should still be interpreted with caution.

Aversion to unconditioned stimuli. Participants were asked to self-report their level of aversion to the unconditioned images and digital sound clips used in the experimental procedures. Four pictures and three sound clips were presented individually in the format described above. After each presentation, participants rated how unpleasant they found the sound clip or image using an on screen Visual Analogue Scale (VAS). Participants submitted a response ranging from 0 to 100 by using the computer mouse to manipulate a slider that moved along a ruler with numerical and descriptive anchors (0 = "neutral", 50 = "unpleasant", 100 = "very unpleasant"). Participant aversion to the visual stimuli was operationalized with the arithmetic mean of the within individual VAS picture ratings (picture VAS). Aversion to the auditory stimuli was operationalized with the mean of within-subject VAS sound clip ratings (clip VAS).

2.3 Procedure

Upon arrival at the laboratory during a designated time slot participants were greeted in the lobby by an experimenter and led to a private room. Participants confirmed their identity and provided informed consent. All completers of the experimental procedures were exposed to every phase of the experiment in the same order. Although individual differences resulted in between subject variation in the total exposure to training and testing trials, the only portion of the experiment that differed systematically among participants was the intervention phase. In this phase participants received the defusion intervention, reappraisal intervention, or control procedure, according to random assignment.

Phase 1: Self-report measures. Participants completed self-report measures including the demographics questionnaire and BIS/BAS scales using the Qualtrics web-based survey platform.

Phase 2: Stimulus equivalence training. The purpose of this phase was to train equivalence classes. Two three-member equivalence classes ([A1, B1, C1] & [A2, B2, C2]) were trained using match to sample in a one-to-many training procedure (see Dymond & Roche, 2009; Dymond et al., 2011, 2008). The following conditional discriminations were trained: A1->B1, A1->C1, A2->B2, A2->C2. On every training trial, a nonsense word (A1 or A2) first appeared in the top center of the computer screen (called the sample stimulus) and was immediately followed by two further nonsense words (B1 and B2 or C1 and C2) positioned in the bottom left and right corners of the screen (called the comparison stimuli). No observing response was required to the sample, and the sample and comparisons remained on-screen until a response was made. Participants made a response by using the mouse to click on one of the two comparison stimuli. Feedback for a correct response was given in the form of a green check mark below the correct choice and a corresponding sound effect. Feedback for an incorrect response was given in the form of a red 'X' mark below the incorrect choice and a corresponding sound effect.

The sequencing of training was modelled after that used by Luciano et al. (2013) and Luciano et al., (2014). Training occurred in three sequential stages, each with a criterion that participants needed to meet before progressing to the next stage: 1) training with the first pair of trial types (A1->B1, A2->B2); 2) training with the second pair of trial types (A1->C1, A2->C2); and 3) training with all trial types.

Stage 1 and 2 each consisted of three blocks. The first block consisted of two trials of one of the relations (e.g. A1->B1). The second block consisted of two trials of the second type of relation (e.g. A2->B2). The third block consisted of four trials, two of each trial type, presented in random order (e.g. A1->B1, A2->B2, A2->B2, A1->B1). Each block was repeated until participants achieved a criterion of 100% correct on that block, at which point they progressed to the next block.

Once participants proceeded through stage 1 and 2 in this fashion, they moved to stage 3 in which they were presented with two blocks of 8 trials each. Each block consisted of two trials of all four relation types in random order. If participants did not achieve the criterion of 100% correct in stage three, they cycled back to stage one.

On meeting the stage 3 training criterion, a single block of 12 trials was presented to test for the emergence of combinatorial mutual entailment (i.e. combined symmetry and transitivity). Each of four trial types (B1->C1, B2-> C2, C1-> B1, C2-> B2) was presented three times in the absence of feedback. Mastery criterion to infer the emergence of stimulus equivalence relations was set at 12 consecutive correct responses. If participants failed to achieve this, they cycled back to stage 1 of the training sequence.

If participants failed to achieve the mastery criterion by the 180th trial (across all training and testing stages), they were thanked for their time and dismissed from later portions of the study. I expected an attrition rate of approximately 5% at this point of the study. The weighted average attrition rate at comparable phases in more recent RFT studies was approximately 0% (Dymond et al., 2011, 2012; Hooper et al., 2010; Luciano et al., 2013, 2014). However, the testing criterion of 100% correct in combination with a preset maximum number of attempts was expected to marginally increase attrition in this phase.

Phase 3: Respondent Conditioning. The purpose of this phase was to train B1 and B2 as a CS^+ and a CS^-, respectively, in order to facilitate operant avoidance training in the proceeding phase (Dymond & Roche, 2009; Luciano et al., 2014). Two trial types were presented in a single block of eight trials (4 of each trial type). In the first trial type, B1 appeared in the center of the screen for 5 s, followed by a 2 s interval, after which either a blank screen was presented or a 600 x 800 pixel aversive photograph and an aversive sound (together the US^+) was presented. The US^+ followed 75% of the presentations of B1 in order to buffer against extinction in later phases of the experiment. In the second trial type, B2 appeared in the center of the screen for 5 s, followed by a 2 s interval, after which a 600 x 800 pixel neutral photograph was presented. The neutral stimuli followed 75% of the presentations of B2. Trials were presented in pseudorandom order with the constraint that no trial type was presented more than twice in a row.

Phase 4: Avoidance Training. The purpose of this phase was to train differential avoidance responding—avoidance in the presence of the B1 stimulus and non-avoidance in the presence of the B2 stimulus. Two operant trial types were presented to participants. In the first trial type, the B1 stimulus was presented, and the US^+ was programmed to follow with a 75% contingency. During the presentation of the B1 stimulus, participants were given the option to make an operant avoidance response by using the mouse to click a button on the screen (marked “Avoid”), which removed the B1 stimulus, cancelled the presentation of the US, and presented the words “consequence cancelled” in the center of the screen. The second trial type followed the same format, except the B2 stimulus was presented as the antecedent and the US^- was programmed as the consequence with a 75% contingency.

The operant trials were conducted in largely the same format as previous studies that utilized an RFT avoidance paradigm (Dymond et al., 2011, 2012; Luciano et al., 2014) with a difference in the method of inputting the avoidance response. Previous studies utilized a key or space bar press, while this study utilized a graphic user interface and the mouse.

In addition to the operant trial types, this phase also included two respondent trial types (see Phase 3), in order to facilitate operant responding and mitigate extinction of the respondent functions in later phases (see Luciano et al, 2014).

This phase consisted of two blocks of 8 trials each: 3 trials each of the 2 operant trial types; and 1 trial each of the two respondent trial types. The trials were presented in pseudorandom order (i.e., no more than two consecutive exposures to each trial type). The two blocks were repeated until participants made 12 consecutive ‘correct’ responses (i.e. avoidance in the presence of B1, and non-avoidance in the presence of B2). If participants did not meet this criterion after 8 blocks, they were thanked for their time and excused from the later portions of the experiment. I did not expect any attrition in this phase of the experiment. The weighted average operant training attrition rate in more recent RFT studies was approximately 1% (Dymond et al., 2011, 2012; Hooper et al., 2010; Luciano et al., 2013, 2014).

Phase 5: Test of transformation of function. This phase tested for the transfer of avoidance responding and non-avoidance responding from B1 to C1 and from B2 to C2, respectively, in largely the same format as the aforementioned RFT studies (Dymond et al., 2011, 2012; Luciano et al., 2014,). The key trials in this phase were operant trials (see Phase 4) with the C1 and C2 stimuli as antecedants. Similar to the B2 operant trials, the C2 stimulus was followed by a US^- 75% of the time. Critically, the C1 stimulus was never followed by an unconditioned stimulus (i.e., there were no aversive or appetitive consequences for non-avoidance). This ensured that the expected avoidance in the presence of the C1 stimuli was due solely to indirect contingencies (the transformation of function according to relational frames; Dymond, Dunsmoor, Vervliet, & Roche, 2015; Dymond & Roche, 2009).

This phase consisted of two blocks of 8 trials each: 3 each of the C1 and C2 operant trials; and 2 each of the B1 and B2 respondent trials (see phase 2). Avoidance in the presence of the C1 stimuli combined with non-avoidance of the C2 stimuli demonstrates the transformation of function and constitutes the criterion of interest. The B1 and B2 respondent trials were interspersed with the C1 and C2 operant trials in order to mitigate extinction of the directly conditioned responses. All trials were presented in a pseudorandom order with the only constraint that no more than two consecutive trials of the same type could occur. Participants were required to exhibit a mastery criterion of 100% avoidance of C1 stimuli and

100% non-avoidance of C2 stimuli (12 trials total). If participant did not meet this criterion, they were cycled back to phase 3 of the study. If participants did not meet criterion a second time, they were thanked for their time and excused from the later portions of the experiment. I expected an attrition rate of approximately 5% at this phase of the study. The weighted average attrition rate of comparable RFT studies at transformation of function phase was approximately 16% (Dymond et al., 2011, 2012; Hooper et al., 2010; Luciano et al., 2013, 2014). I expected a relatively lower attrition rate at this point in the study due to the more stringent training criterion applied in an earlier stage. In addition, the use of arbitrary words, as opposed to abstract symbols or random letters was expected to lower the attrition rate in this phase.

Random assignment. Participants who met all criteria in the preceding phases were randomly assigned to receive one of the three intervention exercises using block randomization.

Phase 6: Intervention or control. In this phase the experimenter delivered one of three experimental manipulations: 1) defusion intervention, 2) cognitive reappraisal intervention 3) control exercise.

Phase 7: Avoidance Testing. This phase served as an outcome test of avoidance behavior in the presence of the indirectly conditioned C1 and C2 stimuli. Participants were presented with the same trials as phase 5 of the experiment, in pseudorandom order, with the addition of a response cost.

Response cost*.* As an analogue to the functionally impairing consequences of overgeneralized avoidance observed in clinical populations, a response cost was added for avoidance responses. Behavioral responses in the presence of competing avoidance and approach contingencies serve as a better indicator of problematic avoidance behavior than behavioral responses in the presences of avoidance contingencies alone (Hayes, Lattal, & Myerson, 1979; Hayes, 1976). Each time an avoidance response was made in the presence of the C1 stimuli, the inter trial interval immediately following the response was increased by a time increment. In addition, the size of the time increment was accreted with every avoidance response. The baseline inter trial interval was set at 5 seconds. The initial time increment for avoidance was set at 1 second and it accreted by one second with each additional avoidance response. Therefore, if a participant avoided all 6 times in the presence of the C1 stimuli, the six inter trial intervals following those six responses would be of length 6, 7, 8, 9, 10, and 11 seconds, respectively.

The increase in inter trial interval could have resulted in groups systematically differing in the amount of time to complete this phase. Therefore, a correlated time delay was added at the end of this phase to offset any systematic difference. The maximum possible cumulative time penalty a participant could have incurred was 21 seconds. The baseline time delay was set at 21 seconds and was adjusted downward depending on the number of avoidance responses made. A participant who avoided zero C1 stimuli experienced a 21 second delay, two C1 stimuli an 18 second delay, three C1 stimuli a 15 second delay, and so forth.

Phase 8: Equivalence testing. In this phase a test of the derived equivalence relations trained in phase 1 was conducted. Participants completed the same testing block completed in phase 1. A single block of 12 trials was presented to test for combinatorial mutual entailment. Each of the four tasks (B1->C1, B2->C2, C1->B1, C2->B2) was presented three times in the absence of feedback. Equivalence responding in line with the trained relations would indicate a stable relational network, while a change in equivalence responding would indicate an effect of the interventions on $C_{rel.}$

Phase 9: **Aversion to unconditioned stimuli**. In this last phase of the study participants completed the picture VAS and sound clip VAS ratings described above.

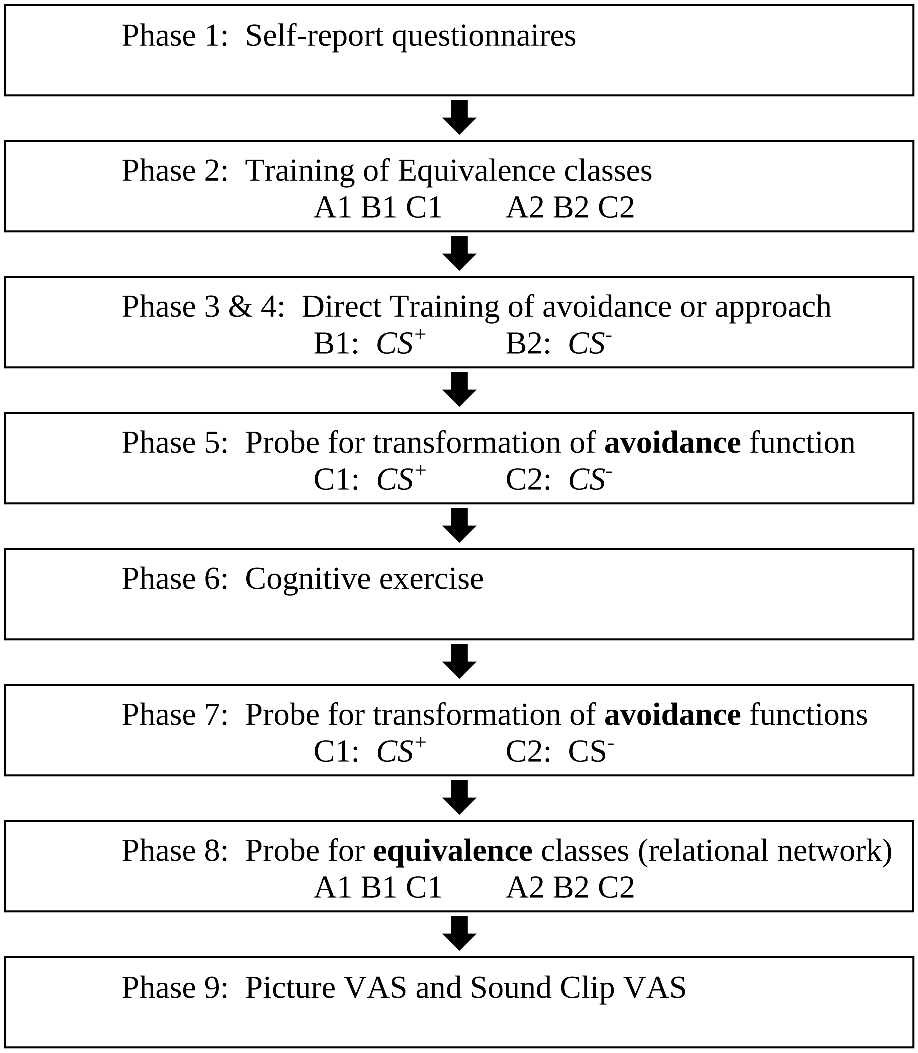

Figure 2.1 Experimental procedure

3 DATA ANALYTIC STRATEGY

3.1 Variables

Outcomes. The two primary outcome datum were post intervention avoidance responding and post intervention equivalence responding. I operationalized avoidance responding with the number of avoidance responses exhibited by each participant in phase seven of the experiment. I operationalized equivalence responding with the number of correct MTS responses exhibited by each participant in phase six of the experiment. Higher numbers corresponded to a more stable or in-tact relational network, and lower numbers correspond to a disruption in the relational network. Similar to the procedures of Luciano et al. (2014) and Roche et al. (2009), equivalence responses were measured after avoidance responses, because of the priority of capturing avoidance responding. However, in the present analyses, these two outcomes are considered as being measured simultaneously.

Potential covariates. In addition to the primary outcome data, I considered several self-reported variables and behavioral indicators for inclusion as potential covariates in the between groups analysis of avoidance responding and the regression analyses. The self-reported variables included baseline levels of behavioral inhibition and behavioral activation and participants' aversion to the unconditioned stimuli. The behavioral indicators estimated participants' ability to learn relational networks and achieve derived transformation of stimulus function.

I tested each potential covariate for between groups differences using one-way independent groups ANOVAs when assumptions of normality and homoscedasticity were met. When the data did not meet assumptions for ANOVA, I utilized the non-parametric alternative, Kruskal-Wallis one-way analysis of variance. I only considered covariates that differed significantly between groups for inclusion in the between groups analysis of avoidance responding.

I also, examined the bivariate relationships between potential covariates and the two main outcomes using correlation coefficients. If there was a significant correlation between a potential covariate and at least one of the outcome measures, I considered it for potential inclusion in the regression analyses.

3.2 Choice of regression model

Avoidance responding. The testing of hypotheses 1 and 2 involved the modelling of avoidance responding data. The avoidance testing phase consisted of a series of six binary decision points—avoid, or do not avoid. I considered two options for summarizing the data. The first option was to summarize the six decision points into a count variable—the number of avoidance responses—which would range from zero to six. The second option was to model avoidance responding as a repeated measures binary variable.

Count regression models. I examined the univariate and conditional distributions of the count of avoidance responses to determine which count regression models would provide the best fit to the data. As is often the case with count variables, the data exhibited many features that made Ordinary Least Squares (OLS) regression unsuitable (Coxe, West, & Aiken, 2009). The univariate distribution was right-skewed, kurtotic, and heteroskedastic, and there were a large proportion of zeroes (43%). Behavioral science statisticians often recommend the Poisson family of regression models (i.e., Poisson, negative binomial, zero-inflated negative binomial) for count outcomes, particularly when the arithmetic mean is below 10 (Coxe et al., 2009).

Since the distribution of avoidance responding was over-dispersed, $\alpha = 0.971$, $p = .003$, I considered a negative binomial model rather than a Poisson regression. Because of the large number of zeroes in the data, I contemplated the zero-inflated negative binomial model. Zero-inflated models assume that the data contain some proportion of structural zeroes. There is a distinct process that results in structural zeroes, as compared to zeroes that had some hypothetical probability of being greater than zero (Coxe et al., 2009). Because I did not have theoretical reasons to suspect structural zeros in the data and all of the participants had demonstrated themselves theoretically capable of producing avoidance responding with some probability in phase 5 of the experiment, I rejected the zero-inflated model, for the negative binomial model.

The negative binomial distribution appeared to fit the data reasonably well with one major exception. In contrast to the unimodal Poisson and negative binomial distributions, the data was multi-modal, with modes at zero and six. An intercept-only negative binomial regression model resulted in a very high standardized residual at the level of six avoidance responses. This appeared to represent a systemic bias to under-estimate the number of participants in this category of responders. This bias persisted as predictors were added to the model.

Because the avoidance distribution was multimodal, I considered a logistic regression model (binary or ordinal), in addition to the negative binomial model. However, the conditional distributions of the ordinal and binary regression models exhibited partial separation, likely due to a small number of participants in one or more of the category combinations (Heinze & Schemper, 2002). For example, none of the participants who exhibited a change in equivalence responding demonstrated any avoidance responding, making this variable perfectly predictive at certain levels. There are several options for adjusting for partial separation, but their power and accuracy are not well established in comparison to other regression models (Heinze & Schemper, 2002), and their use may preclude other data analytic options (e.g. choice of estimator, use of path analysis).

Longitudinal regression models. I considered two types of longitudinal regression models, a repeated measures logistic regression and a survival regression analysis. Visual inspection of the data revealed that once participants exhibited a non-avoidance response, they did not exhibit an avoidance response in later trials. This was likely because after participants withheld an avoidance response the first time, they were able to see that there were no aversive consequences for non-avoidance. Because the first instance of non-avoidance in earlier trials was 100% predictive of avoidance in later trials, a survival regression model was chosen over longitudinal logistic regression. I chose one of the most widely used models, a Cox proportional hazards regression (Kleinbaum & Klein, 2012). The datum predicted by the model was trials until non-avoidance, rather than a count of avoidance responses: a count of zero avoidance responses corresponded to one trial until non-avoidance and five avoidance responses corresponded to six trials until non-avoidance, etc. The parameters of the model estimate the effects that the independent variables have on the hazard (Gelfand, MacKinnon, DeRubeis, & Baraldi, 2016), in this case, the probability of non-avoidance occurring in the next trial. Positive coefficients increase the probability of non-avoidance occurring—reduce avoidance responding—while negative coefficients decrease the probability non-avoidance occurring—increase avoidance responding. The Cox model is a version of the generalized linear model in which the log hazard is predicted as a linear combination of the independent variables. For ease of communication, the terms ***trials until non-avoidance*** and ***avoidance responding*** will be used in place of log hazard in the proceeding analysis and discussion. The Cox model is semi-parametric and has the advantage of requiring fewer distributional assumptions relative to full parametric models such as the ordinal and negative binomial models considered (Kleinbaum & Klein, 2012). In addition, the model accounts for right censoring (Kleinbaum & Klein, 2012). Some proportion of individuals who avoided on every trial would have exhibited non avoidance if there were more trials. Therefore, the avoidance responding variable can be considered right censored, which corresponds with the multiple modes at zero and six. The Cox regression model factors the observed censoring into parameter estimates, while the count regression models considered do not account for right censoring (Kleinbaum & Klein, 2012). The avoidance data met the proportional hazards assumption of the Cox regression model, as determined by time dependent covariate tests.

Given the relative advantages of the Cox proportional hazards regression in modelling the avoidance responding data, I chose it to model hypothesis 2. Due to the relatively small sample size, a robust estimation method was used, maximum likelihood with robust standard errors (MLR; Muthén & Muthén, 2012).

3.3 Analysis of hypotheses

Hypothesis 1: between group analysis of avoidance responding. To test for between groups differences in avoidance, I performed a non-parametric Kruskal-Wallis test and post-hoc paired comparisons utilizing Mann-Whitney tests with a Bonferonni correction.

Hypothesis 2: test for C_{func} effects. To test for moderation of the relationship between avoidance and equivalence responding by treatment group, I fit a Cox proportional hazards regression model to the data. Figure 3.1 details the regression equations used to test moderation, and Figure 3.2 and Figure 3.3 illustrate the hypothesized path model in conceptual and statistical forms, respectively. As can be seen from equation 2 of Figure 3.1, trials until non-avoidance was regressed on the predictor (equivalence responding), the proposed moderator (group membership), and the interaction between the predictor and proposed moderator. A significant contribution to the fit of the model by the interaction terms (b_7 and b_8 in Figure 3.1; $cmod_D$ and $cmod_R$ in Figure 3.3), collectively, would indicate a moderation effect (Hayes, 2013). Since the regression model without the interaction (equation 1 of Figure 3.1) is nested within the moderation model, I performed a likelihood ratio test (Asparouhov, Muthén, & Muthén, 2014) to determine if the interaction made a significant contribution to the likelihood of the model.

Equation 1: $\text{AVOID} = b_1\text{DEF} + b_2\text{REA} + b_3\text{EQU}$

Equation 2: $\text{AVOID} = b_4\text{DEF} + b_5\text{REA} + b_6\text{EQU} + b_7(\text{DEF} \times \text{EQU}) + b_8(\text{REA} \times \text{EQU})$

Figure 3.1 Equations used to test for moderation.

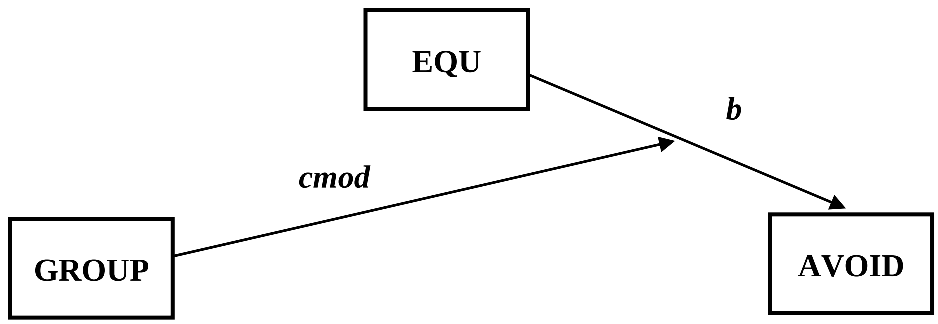

Figure 3.2 Conceptual path diagram of moderation

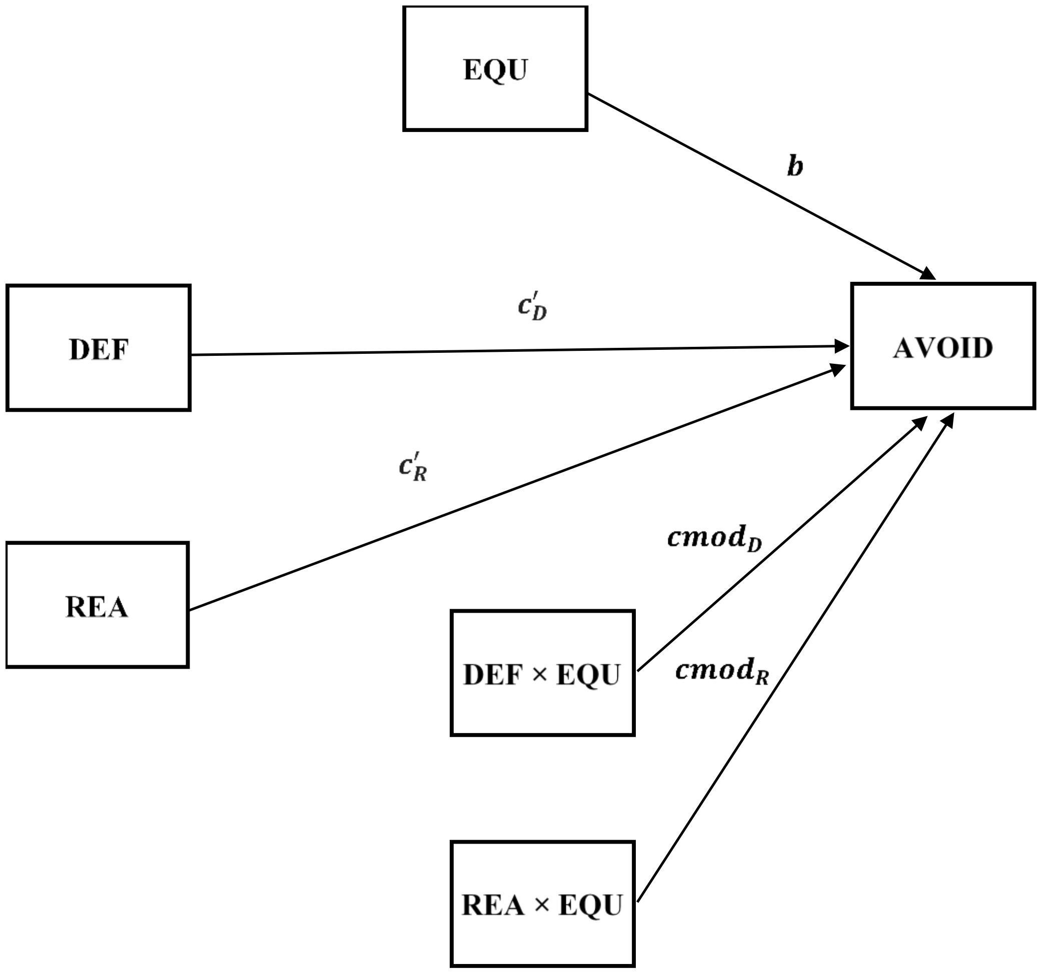

Figure 3.3 Statistical path diagram of moderation.

4 RESULTS

Analyses were performed in SPSS 22, MPLUS 7.4, or SAS 9.4.

4.1 Pre-analysis

Attrition. Ninety-one individuals participated in the laboratory procedures of the study. Of these 91 participants, 82 met criteria for equivalence responding within a pre-specified number of attempts (9.9% attrition). Five of these participants voluntarily withdrew from the study upon beginning phase 2, the respondent conditioning phase (6.1% attrition). Of the 77 participants remaining, 71 achieved operant conditioning criteria within the pre-specified number of attempts (7.8% attrition). Finally, 61 participants demonstrated transformation of stimulus function (phase 6; 14.1% attrition). One participant withdrew from the study upon demonstration of transformation of stimulus function (1.6% attrition). The combined attrition rate was 34.1%. Data analyses presented herein are for the 60 completers of the experimental procedures, except where otherwise noted.

Missing data. There was one case with missing data. Due to experimenter error, one participant did not complete the outcome equivalence responding task (phase 8) and the visual analogue scales (phase 9). The data from this participant was removed on a pairwise basis when calculating between group differences in covariates and bivariate correlation coefficients. Full information maximum likelihood was utilized for all regression analyses (Muthén & Muthén, 2012).

Descriptives and bivariate correlations. For study completers, means and standard deviations for all outcome variables and potential covariates can be found in Table 4.1. Table 4.2 contains a comparison between study completers and non-completers on the BIS/BAS subscale scores. Mann Whitney tests indicated no significant differences between completers and non-completers on the BIS/BAS subscale scores, with *p* values ranging from .48 to .87. Bivariate correlations among all study variables can be found in Table 4.3.

Potential covariates. I considered the four subscales of the BIS/BAS scales as potential covariates: BAS Drive, BAS Fun Seeking, BAS Reward Responsiveness, and BIS. Ability to learn relational networks and ability to achieve transformation of stimulus function were also considered as potential covariates. The number of MTS trials needed to achieve equivalence responding criteria (phase 2) was used to operationalize ability to learn relational networks, and the number of operant trials needed to demonstrate transformation of stimulus function (phase 4 and phase 5) was used to operationalize ability to achieve transformation of stimulus function. Lastly, participants' self-reported aversion to the unconditioned stimuli (picture VAS and clip VAS) were considered as covariates.

In order to test for the assumptions of ANOVA, the distributions of each variable were examined for normality of residuals with visual inspection of residual Q-Q plots and Kolmogorov-Smirnov tests of normality. Heteroscedasticity was also examined using Levene's test. BIS was the only variable to meet the parametric assumptions of normality of residuals and heteroscedasticity. There were no significant differences between groups in any of the potential covariates, and they were not retained in the between groups analysis of avoidance. As can be seen from Table 4.3, the picture visual analogue scale (picture VAS) was the only variable to demonstrate a significant relationship with avoidance responding. I retained this variable as a covariate in the regression analyses of avoidance responding.

4.2 Hypothesis 1

To evaluate differences in avoidance responding between the control group (M = 3.50) reappraisal group (M = 1.35) and defusion group (M = 1.30), I conducted a one-way Kruskal

Wallis analysis of variance. The tested revealed a main effect of group, $\chi^2 = 10.317$, $p = .006$, indicating that one or both of the interventions significantly reduced outcome avoidance responding as compared to the control condition. Post-hoc paired comparisons utilizing Mann-Whitney tests with a Bonferonni correction revealed significant differences between the control group and both the defusion group ($U = 14.525$, $p = .016$) and reappraisal group ($U = 14.575$, $p = .016$). The difference between the two treatment groups was non-significant ($U = 0.050$, $p > .999$).

4.3 Hypothesis 2

To evaluate hypothesis 2, I estimated a Cox proportional hazards regression model. The equation representing the model is presented in Figure 3.1 (equation 2). The outcome variable, trials until non-avoidance, was regressed on the predictor (equivalence responding) the proposed moderator (treatment group; represented by two indicator-coded variables), and the cross-products of the indicator-coded variables and equivalence responding. Table 4.4 contains the regression parameter estimates and Figure 4.1 combines these estimates with a statistical path diagram of the model. A significant contribution to model fit by the cross product terms, collectively, would indicate a significant moderation effect (Hayes, 2013). A likelihood ratio test, with scaling correction for MLR estimation (Asparouhov et al., 2014), indicated that the model with the cross product terms fixed at zero was not significantly less likely than the model with the cross product terms freely estimated ($\chi^2 = 0.585$, $p = .746$). Therefore, the null model, with no interaction between treatment group and equivalence responding, was retained.

4.4 Model re-specification and exploratory analysis

After examination of the data, I specified a larger model, which included and expanded on the analyses proposed for hypothesis 2, on an exploratory basis. Although the original analyses proposed are encompassed within this wider model, looking at the larger variable system made it apparent that I should reformulate hypothesis 2.

Model re-specification. In the CBT theoretical model, interventions target cognitive content, emotions and thoughts, which subsequently leads to behavior change (Hofmann et al., 2013). This framework predicts that the frequency of emotions and thoughts will mediate the effect of cognitive interventions (Hofmann et al., 2013). Translated into the variables of the present study, the CBT model indicates that equivalence responding would mediate the effect of intervention on avoidance responding. Since I hypothesized that the interventions would predominantly function as C_{func} interventions rather than C_{rel} interventions, I excluded the mediation paths from the hypothesized model. The more inclusive mediation model, which encompasses the interaction paths of hypothesis 2 (Hayes, 2013; Kraemer, Kiernan, Essex, & Kupfer, 2008), is depicted in conceptual form in Figure 4.2 and in statistical form in Figure 4.3.

In my a priori model specification, I excluded the *a* path (Figure 4.2) from group to equivalence responding. However, after examination of the data, this appeared to be an error in internal specification. An observed relationship between the reappraisal intervention and equivalence responding indicated a potentially significant *a* path and a possible indirect effect (*ab).* Post hoc adjustments to data models risk over-specifying the models to the chance in the sample observed and inflating experiment-wise type I error rates (Hancock & Mueller, 2010). However, the risk is mitigated in this case because the re-specified model is predicted by a well-established theory (Hancock & Mueller, 2010). Furthermore, modelling best practices suggest that this model should have been specified a priori, as a theoretically plausible competing model (Hancock & Mueller, 2010). Therefore, on an exploratory basis, I tested the full mediation model with direct and indirect effects.

Exploratory hypothesis (reanalysis of hypothesis 2). When referring to the change processes of contextual behavioral interventions, researchers often speak of an expected “desynchrony” between cognitive content and behavior (Hayes et al., 2013, p. 11; Hayes et al., 2011, p. 148-154; Hayes, 2005, p. 137). To test for this desynchrony, several studies of defusion treatments have compared the strength of the association between cognitive content and targeted outcomes (such as behaviors) among treatment groups (Barrera, Szafranski, Ratcliff, Garnaat, & Norton, 2016; Deacon et al., 2011). The expectation was that the defusion treatments would decrease the association between cognitive content and behavior, resulting in an interaction between group and cognitive content in predicting outcomes (Barrera et al., 2016).

Similarly, I hypothesized that the defusion group would moderate the relationship between the content of a relational network and avoidance behavior. There are two potential problems with this approach. One comes from conflating differences in correlation between groups with a statistical interaction. The second is that while desynchrony or moderation may be a part of the overall picture, these results do not adequately capture the hypothesis that defusion and reappraisal work as C_{func} interventions—moderation is a sufficient but not necessary condition to demonstrate a C_{func} intervention.

A difference in association strength between two variables among groups does not necessarily imply an interaction effect (Chaplin, 1991). The former refers to a difference in correlational coefficients, while the latter refers to a difference in slopes—beta coefficients partialed for the unique effects of other covariates (Chaplin, 1991). In the present study, a desynchrony hypothesis would refer to a difference, among groups, in the correlation coefficient between equivalence responding and avoidance. An interaction would involve a difference in the *b* slope coefficient (Figure 3.2) among groups. A moderation analysis does not adequately capture the desynchrony hypothesis; for example, there could be a difference in correlation caused by a difference in variance among groups, without a corresponding interaction (Chaplin, 1991).

Furthermore, while defusion may create a desynchrony between cognitive content and behavior (or an interaction effect), this outcome doesn't necessarily coincide with conceptualization of defusion as a C_{func} manipulation. For example, if the defusion intervention decreased avoidance behavior uniformly across individuals, without altering equivalence responding at all, I would consider this a C_{func} outcome. However, in this scenario, the correlation between equivalence responding and avoidance would not be altered, nor would there be an interaction effect.

In considering all of the ways that defusion could effect avoidance responding—disrupting the relational network (*a* path in Figure 4.2), decreasing avoidance net of any effect on relational responding (*c'* path), altering the impact of the relational network on avoidance (*cmod* path)--the distinction as a C_{func} intervention appears to depend on a reduction in avoidance responding net of any effect on the relational network. This statistical prediction seems to correspond to the hypothesis that defusion reduces behavior *independently* of any effects on frequency of thoughts (Zettle & Hayes, 1986). In the model depicted in Figure 4.2, the "direct effect" of intervention on avoidance, partialed for any effect of relational responding, is the sum of the *c'* and *cmod* paths (Hayes, 2013). Therefore, the hypothesis that defusion and reappraisal are C_{func} interventions translates into a prediction that one or more of these paths will be significant for each of the treatment groups.

Research aim 2: In the treatment of the symbolic generalization of avoidance do defusion and reappraisal function as C_{func} interventions? Is defusion relatively more efficacious in this regard?

Exploratory hypothesis: There will be significant direct effects of the reappraisal and defusion interventions, net of the effect of equivalence responding, on avoidance responding.

Analysis of exploratory hypothesis. A mediation model was fit to the data with trials until non avoidance as the outcome, group membership as the predictor and relational responding as a potential mediator. The moderation paths (Figure 4.3: $cmod_D, cmod_R$) were trimmed from the full model due to the non-significance of these paths (presented above). Trials until non avoidance was modelled with a Cox regression, while the mediator variable, equivalence responding, was modelled using an ordered logistic regression. Parameter estimates were tested for significance, and effect sizes were estimated. The hazards ratios (ψ) corresponding to each of the treatment groups were calculated as indicators of the treatment effect sizes. For each treatment, the hazards ratio is equal to the chance of a participant in the respective group exhibiting a non-avoidance response in the next trial, divided by the chance of a participant in the control group exhibiting a non-avoidance response in the next trial, holding all covariates constant.

The equivalence responding distribution was kurtotic and skewed with a large number of tied observations in a small range. Both OLS and count regression models were unsuitable, but categorical regression was able to accommodate the features of the data. Modelling equivalence responding as an ordered multi-categorical variable with ordered logistic regression had the benefit of preserving variance that would be lost if a binary variable was created (e.g., with a median split) and modeled with a logistic regression. The equivalence responding data met the proportional odds assumption, as indicated by a proportional odds test (χ^2 = 7.383, p = 0.287). The ordered logistic regression is a form of the generalized linear model (Long, 1997). The log odds of a participant moving into a higher category (in this case more equivalence responding) is predicted by a linear combination of independent variables (Long, 1997). Beta coefficients can be interpreted as the change in the log odds of equivalence responding increasing to a higher level, given a one-unit increase in the independent variable. Positive coefficients indicate a positive relationship with equivalence responding and negative coefficients indicate a negative relationship.

I tested for a mediation effect in a path analysis framework utilizing the product of the coefficients method with Monte Carlo confidence intervals. The product of coefficients method has demonstrated a good balance between type I and type II error rates and high accuracy in simulation studies (MacKinnon, Fairchild, & Fritz, 2007). While the magnitude of the direct and indirect effects are difficult to compare when using survival regression analysis, the product of the coefficients method can reliably be used to test for the significance of the indirect effect (VanderWeele, 2011). When simultaneously using logistic regression to model the distribution of the mediator and a Cox regression to model the distribution of the outcome, the sampling distribution of the *ab* parameter (indirect effect) is not known. However, the Monte Carlo method takes as input the *a* and *b* sample statistics and the variances and covariance of these parameters and generates an empirical sampling distribution that can be used to construct confidence intervals (Selig, Preacher, & Little, 2012).

The regression equations constituting the complete model are displayed in Figure 4.4. As illustrated in Figure 4.3, with a multiple categorical predictor, both paths from the predictor to the outcome variable (c'_D, c'_R), and from the predictor to the mediator (α_D, α_R) are estimated separately for each of the treatment groups (Hayes & Preacher, 2014). The result is a set of direct and indirect effects for both the defusion and reappraisal group.

Results of exploratory hypothesis. Table 4.5 contains the parameter estimates resulting from the mediation path analysis, and Figure 4.5 combines a statistical path diagram with the parameter estimates. The indirect effect of reappraisal on trials until non-avoidance by way of equivalence responding was significant, MC 95% CI [0.183, 3.000]. Furthermore, the direct effect of reappraisal on avoidance, accounting for the indirect effect, was non-significant, b = 0.665, p = .163, indicating that the effect of the reappraisal treatment was fully mediated by the change in equivalence responding. The hazard ratio associated with the reappraisal treatment direct effect (ψ = 1.94) indicated that at the end of any given trial, a participant in the reappraisal group was nearly 2 times more likely to exhibit a non-avoidance response in the next trial, than a participant in the control group, holding all other covariates constant.

In contrast to the indirect effect of reappraisal, the indirect effect of defusion on avoidance responding by way of equivalence responding was non-significant, MC 95% CI [-.0191, 2.492], indicating that the effect of defusion was not significantly mediated by equivalence responding. In addition, the direct effect of defusion on outcome avoidance, after accounting for the potential mediator, remained significant, b = 0.955, p = 0.020. The hazard ratio associated with the defusion direct effect (ψ = 2.60) indicated that at the end of any given trial a participant in the defusion group was 2.6 times more likely to exhibit a non-avoidance response in the next trial than a participant in the control group, all else being equal. This corresponds to a 72% percent chance that a participant treated with defusion would stop avoiding sooner than a participant who received the control exercise, all else being equal

The implication is that the defusion treatment successfully altered the transformation of stimulus function, without a corresponding alteration in the relational network. In contrast, the reappraisal intervention appears to have relied more on a weakening of the relational network to mediate the effect of the intervention.

The size of the reappraisal direct effect (ψ = 1.96) appears to be clinically meaningful, despite the non-significant hypothesis test. A post-hoc power analysis revealed that this study achieved a power (1-β) to detect a reappraisal direct effect of approximately 39% to 54%, depending on the assumption of overlapping variance between the reappraisal treatment and the other covariates (0% to 33.5%). Therefore, the non-significance of the reappraisal direct effect may have been due to a lack of power.

Notwithstanding this possibility, the overall pattern of results—differences between the treatment conditions in the size of indirect effects and direct effects—supports a difference in mechanisms between the two treatment conditions, with defusion indicated more as a C_{func} intervention and reappraisal as a C_{rel} intervention.

Table 4.1 Means and standard deviations for potential covariates and primary outcome measures, for overall sample and by condition

	Overall		Control		Reappraisal		Defusion	
Variable	Mean	*SD*	Mean	*SD*	Mean	*SD*	Mean	*SD*
Avoidance	2.05	2.49	3.50	2.61	1.35	2.16	1.30	2.13
Equivalence responding	11.46	1.04	11.85	0.67	11.00	1.33	11.50	0.89
Trials to equivalence	82.43	38.63	75.70	32.96	87.80	41.16	83.80	42.11
Trials to transformation	50.72	17.09	49.75	15.66	52.60	14.56	49.80	21.07
BAS drive	12.07	2.18	11.35	2.56	12.45	2.33	12.40	1.39
BAS fun seeking	12.77	2.09	12.35	1.73	13.15	1.81	12.80	2.65
BAS reward responsiveness	18.28	1.79	18.35	1.93	17.80	1.77	18.70	1.63
BIS	21.32	3.40	22.40	3.36	20.05	3.30	21.50	3.27
Sound clip VAS	71.85	20.71	71.32	25.75	74.04	19.88	70.32	16.24
Picture VAS	84.98	20.99	83.21	24.59	87.75	19.69	84.11	19.01

Note. Avoidance = count of outcome avoidance responses; Trials to equivalence = number of MTS trials needed to reach pre-intervention equivalence criterion; Trials to transformation = number of operant trials needed to reach pre-intervention transformation of function criterion. Sound Clip VAS = within subject mean of visual analogue scale ratings of aversion to sound clips; Picture VAS = within subject mean of visual analogue scale ratings of aversion to pictures

Table 4.2 Means and standard deviations for BIS/BAS scales scores by study completion status

	Completers		Non-Completers	
Variable	Mean	*SD*	Mean	*SD*
BAS drive	12.07	2.18	11.77	1.98
BAS fun seeking	12.77	2.09	12.65	1.91
BAS reward responsiveness	18.28	1.79	18.23	1.86
BIS	21.32	3.40	21.45	3.91

Table 4.3 Bivariate correlations among primary outcome measures and potential covariates

Variable	1	2	3	4	5	6	7	8	9
Avoidance									
Equivalence responding	0.351**								
Trials to equivalence	-0.068	-0.084							
Trials to transformation	-0.103	-0.106	-0.062						
BAS drive	-0.049	-0.199	0.282**	0.071					
BAS fun seeking	0.045	-0.195	0.023	0.053	0.398				
BAS reward responsiveness	0.066	-0.089	0.081	0.023	0.441**	0.241			
BIS	0.113	0.162	-0.027	0.042	-0.170	-0.258**	0.086		
Sound Clip VAS	0.150	-0.042	-0.152	0.212*	-0.020	0.009	-0.089	0.092	
Picture VAS	0.283*	0.057	-0.052	0.194	0.062	-0.004	-0.022	0.110	0.432

Note. Avoidance = count of outcome avoidance responses; Trials to equivalence = number of MTS trials needed to reach pre-intervention equivalence criterion; Trials to transformation = number of operant trials needed to reach pre-intervention transformation of function criterion. Sound Clip VAS = within subject mean of visual analogue scale ratings of aversion to sound clips; Picture VAS = within subject mean of visual analogue scale ratings of aversion to pictures. Correlation coefficients between picture VAS and avoidance is a Pearson product moment correlation, while all other coefficients are Kendall rank correlations.
*$p < .05$. **$p < .01$

Table 4.4 Summary of moderation regression analyses

Variable or statistic	*B*	*SE*
Equation 1		
PIC	-0.020**	0.006
Defusion (DEF)	0.955*	0.390
Reappraisal (REA)	0.665	0.390
Equivalence Responding (EQU)	-0.564***	0.080
Equation 2		
PIC	-0.020**	0.006
DEF	1.217	2.283
REA	-0.744	2.131
EQU	-0.602***	0.099
DEF × EQU	-0.023	0.218
REA × EQU	0.128	0.208

Note. PIC = within subject mean of visual analogue scale ratings of aversion to pictures
$*p < .05$. $**p < .01$. $***p < .001$.

Table 4.5 Summary of mediation regression analyses

Variable or statistic	*b*	*SE*	*ψ*
Equation 1, dependent variable: equivalence responding (EQU)			
Defusion (DEF)	-1.960	1.164	
Reappraisal (REA)	-2.677*	1.190	
Equation 2, dependent variable: trials until non-avoidance (AVOID)			
PIC	-0.020**	0.006	0.980
DEF	0.955*	0.412	2.599
REA	0.665	0.476	1.944
EQU	-0.564***	0.141	0.569

Note. PIC = within subject mean of visual analogue scale ratings of aversion to pictures. *ψ* = hazards ratio, a measure of effect size.
*$p < .05$. **$p < .01$. ***$p < .001$.

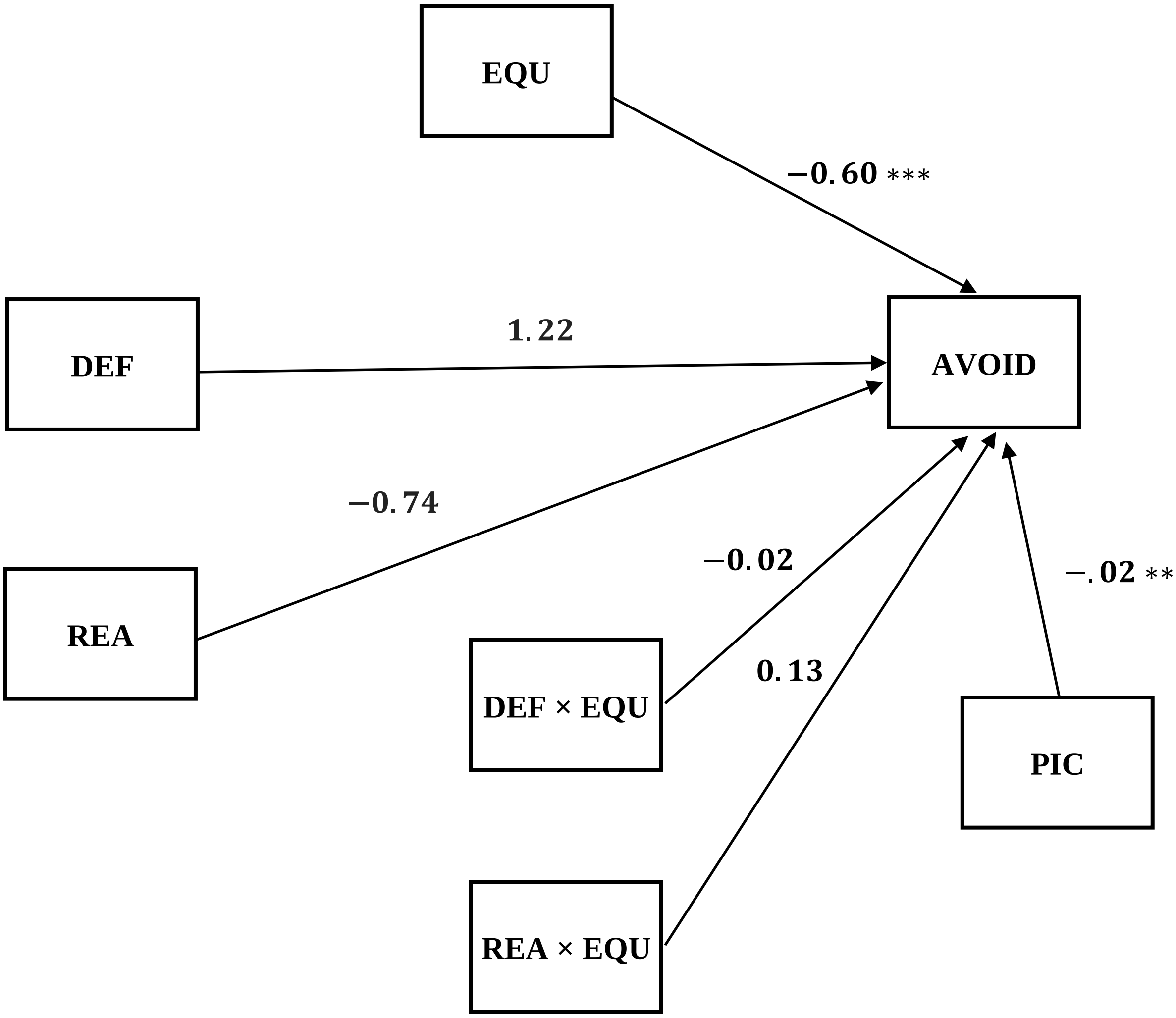

Figure 4.1 Statistical path diagram of moderation with path coefficients and covariate

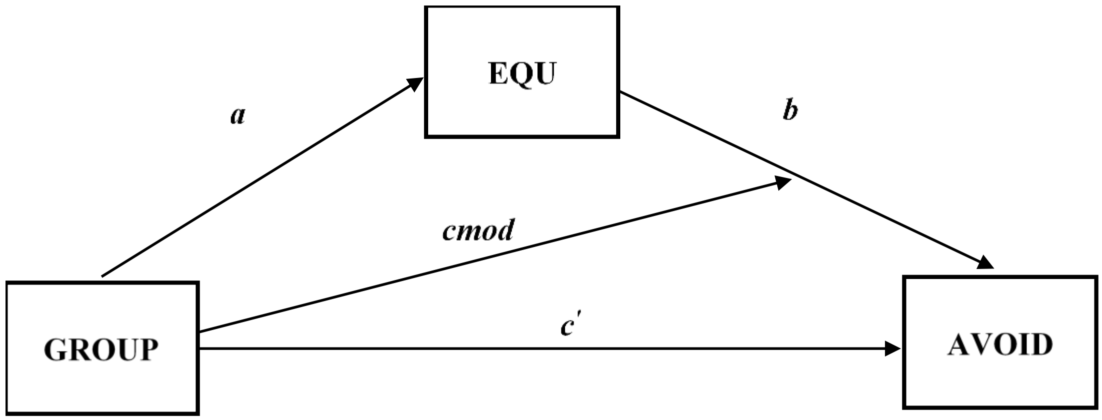

Figure 4.2 Conceptual path diagram of mediation including interaction between mediator and predictor

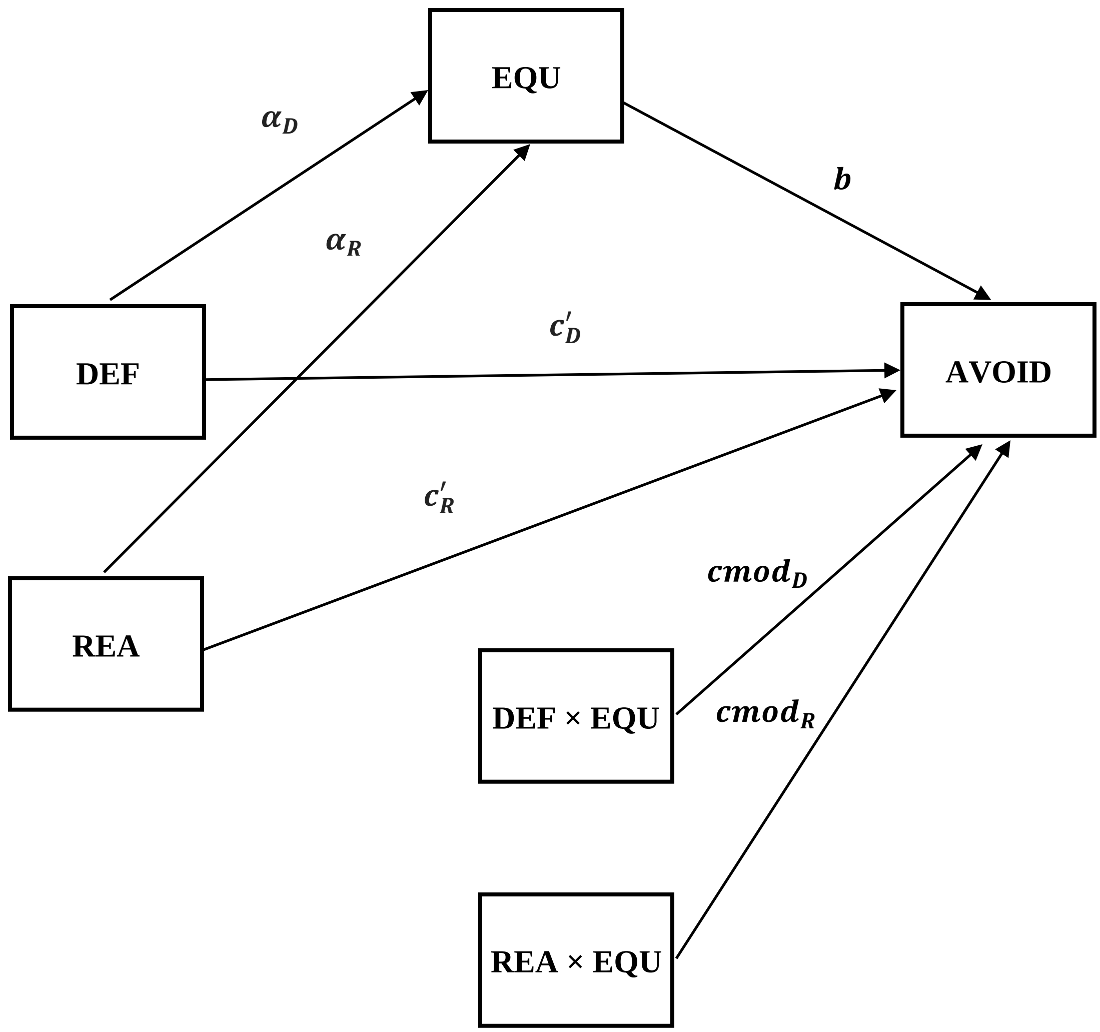

Figure 4.3 Statistical path diagram of mediation

Equation 1: $\mathrm{EQU} = \mathrm{b}_9\mathrm{DEF} + \mathrm{b}_{11}\mathrm{REA}$

Equation 2: $\mathrm{AVOID} = \mathrm{b}_4\mathrm{PIC} + \mathrm{b}_5\mathrm{DEF} + \mathrm{b}_6\mathrm{REA} + \mathrm{b}_7\mathrm{REL}$

Figure 4.4 Equations used to test for mediation.

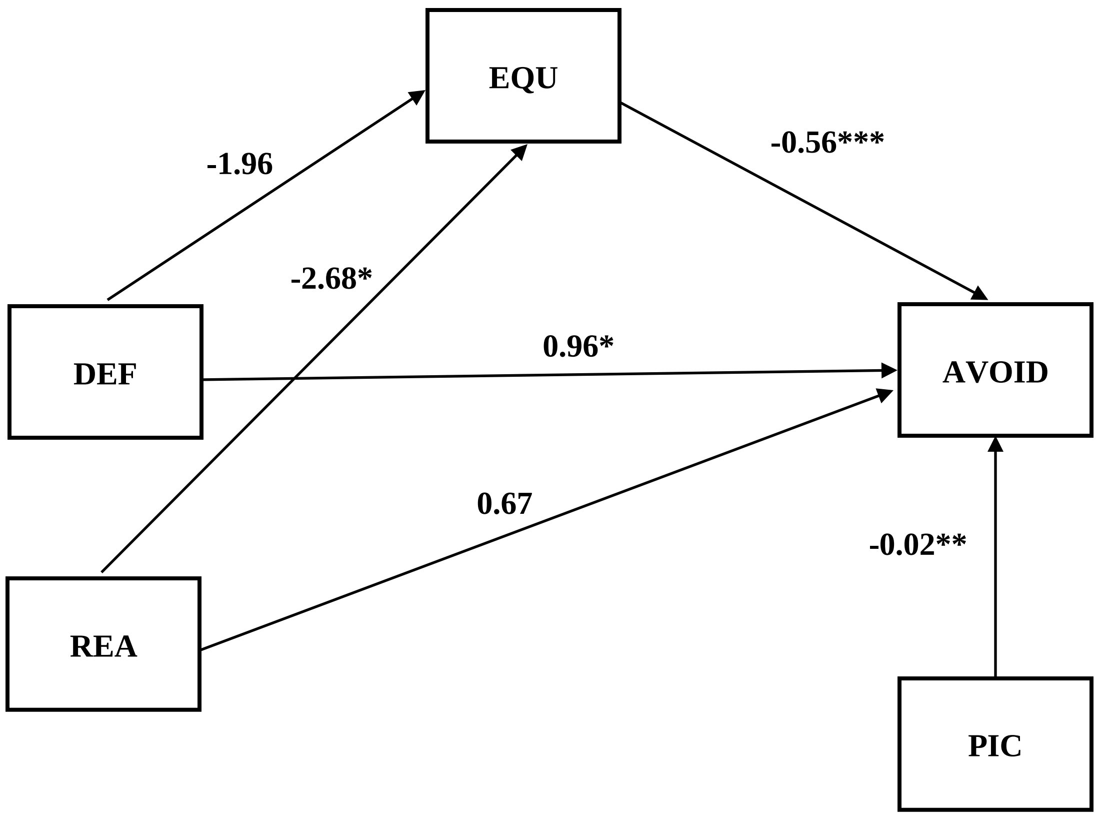

Figure 4.5 Statistical path diagram of mediation including path coefficients

5 DISCUSSION

CBS philosophy and research agenda. In recent years there has been a call for efforts to bolster the connection between underlying theory and treatment technology in CBT (Dobson & Beshai, 2013; Herbert et al., 2013). CBS, largely consisting of the research agendas of RFT and ACT, represents one branch of that effort, with its aspiration to fully integrate an explicit philosophy of science, a body of theory, and treatment techniques (Hayes, Barnes-Holmes, et al., 2012; Vilardaga, Hayes, Levin, & Muto, 2009). Departing from the prevailing mechanical realism of most CBT theory, the CBS scientific philosophy of functional contextualism dictates that the veracity of theoretical principles be determined by their contribution to the goal of prediction and influence (Hayes, Barnes-Holmes, et al., 2012). The application of this philosophy of science to psychology has had perhaps the greatest impact in the areas of the theoretical modeling and clinical treatment of cognition (Barrera et al., 2016). The theoretical supposition that the functional impact of thinking can be controlled separately from the content of thinking may be the single greatest RFT contribution to date (Hayes, 2014, August 11), and the corresponding applied construct, cognitive defusion, may be the single biggest distinction between ACT and traditional CBTs (Barrera et al., 2016; Herbert & Forman, 2013; Hofmann et al., 2013).

Defusion construct. Notwithstanding these developments, some researchers question whether the CBS research program has resulted in applied interventions that are fundamentally distinct from that of CBT (Arch & Craske, 2008; Hofmann & Asmundson, 2008). Not surprisingly, much of this discussion has focused on the treatment of cognition, generally, and comparisons between cognitive defusion and cognitive change strategies, specifically. CBS has encouraged research on mechanisms and treatment components from very early in the development of ACT (Hayes et al., 2006; Levin, Hildebrandt, Lillis, & Hayes, 2012; Vilardaga, Hayes, Levin, & Muto, 2009). However, the evidence resulting from this effort is mixed. While several laboratory studies have demonstrated results consistent with the theory of defusion, several have not, and these studies are generally limited by reliance on self-report evidence and wide variation in the operationalization and measurement of the defusion construct (Deacon et al., 2011; Levin et al., 2012; Masuda et al., 2010, 2004; Yovel et al., 2014). The state of the evidence base has led some to posit that defusion and cognitive change strategies (particularly reappraisal) are indistinguishable and work according to the same mechanisms (Mennin et al., 2013).

Theory validation. As argued in detail above, it appears that the current limitations of the research base surrounding defusion and related cognitive change strategies derive from the lack of empirically validated principles based definitions of the constructs. RFT researchers have offered a general principles based definition of cognitive interventions as contextual manipulations of relational framing, with a distinction between strategies that rely on changing thought content to change behavior (C_{rel} strategies), and those that change behavior without targeting thought content (C_{func} strategies) (Barnes-Holmes et al., 2001; Hayes et al., 2013). The RFT definition appears to be parsimonious, internally coherent, and consistent with the applied theories of both CBT and ACT. Still, theoretical principles should be validated by empirically determining their relevance to intervention techniques (Herbert et al., 2013; Levin et al., 2012). This step has yet to be taken by CBS researchers, even though the methods for doing so have already been proposed (Blackledge, 2007; Luciano et al., 2014; Roche et al., 2009). As a result, the distinction between C_{rel} and C_{func} strategies continues to be strongly debated even within the CBS community (Hayes, 2014; Villatte, 2014).

Present study. This study represents an effort to elucidate the connection between RFT and the interventions of defusion and reappraisal by examining the effects of exemplars of these techniques on relational framing behavior (indirectly acquired maladaptive avoidance.) The expected reductions in avoidance behavior would demonstrate these constructs as manipulations of the context of relational framing. In addition, reductions in avoidance without a corresponding disruption in the relational networks would demonstrate these interventions as C_{func} manipulations. Alternatively, if changes in avoidance relied on disruption of the relational network, this would establish them as C_{rel} interventions. Consistent with the theoretical underpinning of defusion, I hypothesized that the commonly used defusion strategy, rapid vocal repetition, would function as a C_{func} intervention. Given the parity in efficacy between defusion and cognitive change strategies (Deacon et al., 2011; Yovel et al., 2014), the preliminary evidence that CBT may work according to defusion mechanisms (Arch et al., 2012), and the conceptual similarity between defusion and reappraisal (Mennin et al., 2013), I expected that reappraisal would function similarly to defusion as a C_{func} intervention.

5.1 Findings

Reductions in avoidance behavior. Consistent with my hypotheses, the present defusion and reappraisal strategies significantly reduced indirectly acquired maladaptive avoidance responding, demonstrating evidence for these interventions as manipulations of relational framing behavior. Given that the present defusion strategy was designed to target a more malleable aspect of relational framing, the transformation of function, I expected that defusion would be relatively more efficacious in reducing avoidance. However, the results did not support this hypothesis. Both the defusion strategy and the reappraisal strategy worked similarly well in reducing avoidance behavior.

C_{func} vs. C_{rel}. At the level of mechanisms, I expected the present reappraisal and defusion strategies to function similarly as C_{func} interventions, with defusion demonstrating a relative advantage in this regard. I proposed that a desynchrony between relating behavior and avoidance behavior would demonstrate these interventions as C_{func} manipulations, and I tested for this desynchrony with a moderation analysis. However, the moderation effects were non-significant. The null findings may be due to the limited variability in equivalence responding observed in this study—an artifact largely resulting from the need for participants to achieve a set criterion in equivalence responding before receiving the interventions. Every participant started at the same level of equivalence responding prior to intervention, and since a criterion of 100% correct was required, the variable could only be altered in one direction.

Notwithstanding the possibility of statistical artifacts, there was a discrepancy between the null findings and the fact that the defusion intervention ostensibly functioned independently of equivalence responding, i.e. the relationship between the defusion intervention and equivalence responding was non-significant. This discrepancy led to a post hoc consideration of the conditions that are necessary and sufficient to demonstrate interventions as C_{func} manipulations. I concluded that a unique effect of intervention, partialed for the effect of equivalence responding, was necessary to demonstrate a disruption in the transformation of function. This direct effect could involve a moderation of the effect of equivalence responding on avoidance, a unique effect of intervention partialed for the effect of relating, or both.

After observing that the reappraisal intervention apparently weakened the experimental relational networks, I concluded that mediation of the intervention effects by changes in equivalence responding, consistent with CBT theory, should be examined along with the hypothesized direct effects.

Overall, an analysis of the direct and indirect paths from the interventions to avoidance behavior supported a distinction between defusion as a C_{func} intervention and reappraisal as a C_{rel} intervention. The effect of the defusion intervention was due to a direct effect on avoidance. In contrast, the effect of reappraisal was fully mediated by equivalence responding, rendering the direct effect non-significant.

5.2 Conceptual and applied implications

State of the research. Cognitive defusion, was first proposed as a central treatment mechanism over thirty years ago (Zettle & Hayes, 1986; Zettle, 2011). The construct arose from the behavioral analysis of client-therapist interactions and was justified with appeals to the basic analysis of verbal behavior and the principles of RFT (Blackledge, 2007; Zettle & Hayes, 1986). Despite the behavioral origins of the defusion construct and the CBS priority placed on depth in research, this is one of only a few studies to examine defusion using behaviorally informed outcomes (Hinton & Gaynor, 2008; Hooper & McHugh, 2013; Kishita, Muto, Ohtsuki, & Barnes-Holmes, 2014). Furthermore, this is the first study to test the RFT mechanisms of a standalone defusion intervention and the first study to compare those mechanisms to a cognitive change strategy.

Applied research implications. Viewed from the applied level of analysis, this study furthers the research effort of using laboratory-based studies to examine the theoretical models of ACT and CBT treatment components. Though laboratory based methods should not be considered a substitute for the testing of clinical efficacy in component and treatment studies, their level of precision and control allows for the validation of theory and testing of principles (David & Montgomery, 2011; Hayes et al., 2013; Kazdin, 1978; Levin et al., 2012). By isolating treatment components, laboratory studies have demonstrated that defusion and reappraisal techniques work to reduce self-reported distress and that defusion works to alter self-reported metacognition (Gross, 1998, 2002; Levin et al., 2012; Masuda et al., 2004). However, distress is neither a treatment target nor theorized mechanism of action in ACT (Hayes, et al., 2012). Although researchers posit that metacognition is related to the theoretical principles of defusion techniques, previous studies have not been able to establish this connection, nor have they been able to establish the relevance of self-reported metacognition results to the target of treatment, values-based behavior (Levin et al, 2012). The results of my thesis demonstrated that the particular defusion and reappraisal interventions employed worked as expected to reduce symbolically generalized maladaptive avoidance behavior—a class of behavior which often functions to impede or reduce value-directed action in clinical populations (Chawla & Ostafin, 2007; Kashdan, Barrios, Forsyth, & Steger, 2006). Therefore, this study extends the existing laboratory evidence by demonstrating a connection between the treatment components and outcomes that are directly relevant to the treatment goal of value-based behavior (Forsyth et al., 2006; Levin et al., 2012).

Construct validation. This study also provides evidence for the construct of defusion as defined at the applied level by the ACT model (Hayes et al., 2013) and as defined at the basic level with behavioral principles (Blackledge, 2007; Snyder, Lambert, & Twohig, 2011). The aforementioned self-report evidence provided preliminary support that defusion works as theorized to impact the function of thoughts (Deacon et al., 2011; Masuda et al., 2010; Masuda et al., 2004; Masuda et al., 2009; Yovel et al., 2014). However, it is debatable how well these self-report observations capture the behavioral principles involved (e.g., capture the function of verbal behavior) and how relevant they are to the prediction and influence of verbal behavior (Blackledge, 2007; Levin et al., 2012; Masuda et al., 2004). The present study helped to answer these questions by directly and precisely measuring the verbal behavioral principles of defusion outcomes, and demonstrating that, in parallel to the self-report evidence, defusion alters function independently of the form of verbal behavior.

Additional construct validity evidence comes from the observed differences in mechanisms between defusion and the boundary construct of reappraisal. Although not hypothesized a priori, the effect of reappraisal was found to be fully mediated by changes in the relational network (C_{rel}). This set of finding supports the supposition that cognitive defusion is a construct distinct from traditional cognitive change strategies.

Modelling C_{func} effects. The data analyses performed in the present study suggest that the defusion effect is best modelled as an independent (independent of the relational network) effect on behavior. Although, the defusion effect is often spoken about in terms of the disruption of the thought-behavior link or as creating a desynchrony between behavior and private events (Hayes, 2005; Hayes et al., 2011), it may be more accurate to speak about defusion as altering function independently of the form or frequency thinking. Firstly, RFT theory does not appear to support a complete disruption of the thought-behavior link. According to RFT, if relating were to cease there would be no basis for the transformation of function. Said another way, if the frequency of a problematic thought were to decline, the behavioral impact of that thinking would also decline. Secondly, as reasoned earlier and evidenced by the results of this study, moderation is a sufficient but not necessary condition to demonstrate C_{func} effects. Therefore, instead of speaking to a desynchrony, it would be more accurate to speak to an effect independent of the relational networks involved, at least for the present defusion strategy (rapid vocal repetition).

Basic research implications. Viewed from the basic behavioral level of analysis, this study provides support for the RFT model of cognition and cognitive interventions. Together with Luciano et al., 2014, the results demonstrate the suitability of RFT to model and predict both defusion techniques and cognitive change strategies. The conceptualization of cognitive interventions as manipulations of the context of relational framing and the distinction between C_{func} and C_{rel} manipulations appear to be useful for increasing the explanatory power of the theoretical models of cognition (Dymond et al., 2015).

Clinical implications. From the perspective of potential clinical implications, the distinctions between cognitive defusion and cognitive reappraisal observed in this study appear to be particularly relevant. When considered alongside the learning principles of extinction and verbal behavior, they imply that cognitive change strategies should be less effective than cognitive defusion (Hayes et al., 2013, 2006; Hayes, Strosahl, et al., 2012). This is consistent with the ACT model and represents one of the major justifications for the use of ACT over CBT (Barrera et al., 2016; Hayes et al., 2013, 2006; Hayes, Strosahl, et al., 2012). Yet, this implication has not proven out empirically. The present study found no evidence for differences in efficacy between defusion and reappraisal in targeting derived avoidance-evoking functions. This is consistent with the majority of studies comparing defusion and other cognitive change strategies, primarily cognitive restructuring (Deacon et al., 2011; Pilecki & Mckay, 2012; Yovel et al., 2014). Moreover, these findings are largely paralleled at the level of overall treatment: although ACT appears to operate through theoretically consistent mechanisms, the results of randomized controlled trials do not support a difference in efficacy between ACT and CBT (for a recent meta-analysis see Öst, 2014).

5.3 Resolutions

Implied inferiority of C_{rel} manipulations. In the present study, both defusion and reappraisal interventions worked similarly well. This finding disputes the assumption that relational interventions (e.g., content-focused cognitive strategies) are inferior to functional context interventions (e.g., defusion strategies). This discrepancy appears to call into question the logical extensions derived from the RFT application of behavioral principles to cognitive interventions. Those logical extensions include the suppositions that modifying relations is problematic because (1) there are a vast number of environmental stimuli that can evoke any given relation; (2) there is a tendency for punished relations to re-establish in new contexts; and (3) relational framing is practically self-reinforcing because of the ubiquitous context of sense-making (Hayes et al., 2006; Hayes, 2004; Hayes et al., 2013). The exploratory results of this study could indicate that there are problems with one or more of these extensions.

Alternatively, the discrepancy between the exploratory findings and the logical extensions may be due to the particular type of relational responding manipulated in this study. The present study conducted a post-test of avoidance and equivalence responding at only one time point and in only one context, immediately after the interventions were delivered. However, the factors assumed to work against C_{rel} interventions may not come into play until more time has elapsed and individuals are exposed to additional contexts or have contacted additional contingencies of sense-making. Moreover, the length of learning history and complexity of problematic relations may increase the number of outside relations and contexts that could function as antecedents for the problematic relations. Longer learning histories, more complex relations, and longitudinal assessments may yet reveal differences in effectiveness between relational context and functional context interventions.

Attrition. The attrition rate of 34% was considerably higher than the expected attrition rate of 10.3% and should be addressed. The level of attrition relative to other RFT studies differed depending on the stage of the study. That is, the attrition at the stage of testing the derived transformation of stimulus function was comparable to other studies (Dymond, Roche, Forsyth, Whelan, & Rhoden, 2007; Dymond et al., 2008; Luciano et al., 2013, 2014; Roche et al., 2008). However, the attrition at the relational training and operant training stages of the study was considerably higher than other studies in the field (Dymond et al., 2007, 2008; Luciano et al., 2013, 2014; Roche et al., 2008).

A number of factors may have contributed to the higher attrition rate at the equivalence responding testing stage of the study. The current study imposed a preset maximum number of cycles through training and testing, unlike several studies that allowed participants to cycle through until achieving criteria (Dymond et al., 2011, 2012; Hooper et al., 2010; Luciano et al., 2013, 2014). In addition, the criterion of 100% correct on testing trials was more stringent than that used in several studies (Dymond et al., 2011, 2012; Hooper et al., 2010). Lastly, samples could have differed in terms of participant motivation and effort due to participant compensation methods. Unlike the present study, the majority of studies surveyed offered participants monetary compensation (Dymond et al., 2011, 2012; Luciano et al., 2013, 2014).

A large portion of the attrition during the operant testing portion of the study was due to participants' voluntarily withdrawing from the study. In comparison, this source of attrition was not reported in the studies surveyed (Dymond et al., 2011, 2007, 2008, 2012; Luciano et al., 2013, 2014; Roche et al., 2008). The attrition rate during operant testing without voluntary withdrawal was approximately half of that observed.

Generalizability. A claim could be made that the results of this study are not generalizable due to characteristics of the sampling procedures (i.e., non-clinical, mostly female, college student sample) and the laboratory setting (i.e., not ecologically valid). However, the generalizability or external validity of any study should only be judged in the context of the study's purpose, with consideration of what precisely is being generalized (Berkowitz & Donnerstein, 1982; Mook, 1983). If the purpose of this study was to generalize base rates observed in the sample (e.g., the percentage of people who respond to a defusion intervention), or surface characteristics of the laboratory procedures (e.g., aversive photographs resemble feared consequences in the real world), then the question of external validity, as traditionally defined, would be central to the study's claims (Berkowitz & Donnerstein, 1982). However, the purpose of this reverse-translational experiment was to test theoretical relationships among functionally defined behaviors and constructs. Therefore, the generalizability of the results depends largely on two factors: 1) the scope of the theoretical principles demonstrated; and 2) how representative each variable is of the functional class to which it belongs (e.g., how representative is aversive photo presentation of the functional class, punishing consequences; Hayes & Berens, 2004; Mook, 1983). Both factors represent empirical questions which cannot be determined by examining the surface characteristics of the study procedures (Berkowitz & Donnerstein, 1982; Mook, 1983). Furthermore, while future generalizability evidence may qualify or specify the theoretical principles in question, it will not invalidate this study's main contribution as a demonstration of theoretically specified functional relationships between directly observed behaviors.

5.4 Limitations

Mediator assessment. Though the present study provided useful data regarding the behavioral analysis of cognition and cognitive interventions, there were several methodological choices that qualify the conclusions that can be made. Firstly, the potential mediator, equivalence responding, was measured after the outcome variable, avoidance, which limits my ability to determine the directionality of the relationship and make causal inference (Hayes, 2013; MacKinnon et al., 2007). It is feasible that the direction of influence could be opposite that assumed. For example, if participants' non-avoidance led them to directly observe that the contingences for responding differed between the C1 and B1 stimuli, this could have led them to conclude that the two stimuli should not be placed in a relation of coordination. The fact that equivalence responding was not associated with level of avoidance in the control condition suggests that non-avoidance did not cause changes in relating. Yet, the measurement of each variable at multiple time points would better suit the purpose of establishing temporal precedence of the proposed mediator (MacKinnon et al., 2007).

Non-active control. Another limitation of the study involved the use of a non-active control condition. The use of a non-active control could have led to differences in demand characteristics between the control and intervention conditions. For example, both of the interventions targeted the C1 stimulus and spoke to the stimulus directly and the control did not. It is possible that just mentioning the stimulus led individuals to scrutinize their choices regarding the stimulus. Follow-up studies could control for this possibility by implementing an active control condition that specifically targets the C1 stimulus in some fashion.

Participant rule formation. Furthermore, future studies may consider controlling for participant learning, extraneous to the trained relations of coordination, by using self-report assessments regarding participants' rule formation. For example, Dymond et al. (2011) used self-reported expectation ratings to determine if participants' avoidance responses coincided with expectation of aversive contingencies. A post experiment interview could also be used to explicitly assess verbal rules.

Aversion to response cost. In order to statistically control for potential covariates, the present study measured general behavioral inhibition and activation and aversion to the unconditioned stimuli. However, level of aversion to the response cost, increasing time between trials, was not measured. It is possible that statistical precision of the main study variables could have been increased with the inclusion of response cost aversion in the data analytic models.

Sample size. Lastly, the relatively small sample size of this study limited my ability to statistically control for variables and still maintain adequate precision of parameter estimates. Furthermore, a larger sample size could have also resulted in a statistically significant mediation effect in the defusion condition.

5.5 Future directions

RFT modelling of intervention delivery. The findings of this study warrant further reverse translational research efforts to develop finer grained principles-based definitions of cognitive interventions. I anticipate that in future studies, the interventions could be delivered in the same format as the relational networks—training with computer learning tasks—in order to model the specific relations and behavioral operants involved in delivering the interventions. As an illustration, consider the defusion intervention, rapid word repetition. One possibility is that the violation of language parameters acts as an establishing operation (Wray, Freund, & Dougher, 2009) which creates a context of discrimination, as opposed to a context of generalization, between the targeted word and its referents (Dymond et al., 2015). More fine grained principles based modelling of interventions along these lines could help focus and refine existing interventions and help design future interventions.

Forms of problematic relating. Another area of future research involves the manipulation of the types of relations studied. As exhibited earlier during the discussion of the efficacy of C_{rel} interventions, the type of relation studied, in terms of length of learning history, complexity, and longitudinal course, may determine what types of interventions are expected to be the most effective (Blackledge, 2007; Dymond, Ng, & Whelan, 2013; Valverde, Luciano, & Barnes-Holmes, 2009). Moreover, modifying these parameters may provide more accurate models of the classes of verbal behavior that lead to clinical distress and functional impairment at the applied level. Future studies should strive to delineate the parameters and boundary conditions under which there are clinically meaningful differences in how problematic relations function, in terms of expected course and response to different types of intervention.

5.6 Conclusion

In a tightly controlled environment where participants' learning histories with the experimental stimuli could be determined with precision and outcomes could be directly observed, this study demonstrated that commonly used cognitive defusion and cognitive reappraisal strategies are able to effectively target indirectly learned maladaptive avoidance behavior. In addition, it provided construct validity evidence for the theoretical models of defusion at the levels of both ACT theory and RFT. The exploratory results provided initial support for the RFT distinction between defusion, as a manipulation of the functional context of symbolic relating, and reappraisal, as a manipulation of the relational context of symbolic relating—upholding the primary distinction between contextual behavioral and cognitive behavioral treatments. Future reverse-translational work, which further operationalizes applied constructs in terms of RFT behavioral principles, is warranted. Promising lines of research involve behavioral implementation of cognitive interventions and the parametric variation of relational framing in the laboratory.

REFERENCES

Ahmed, O., & Lovibond, P. F. (2014). The impact of previously learned feature-relevance on generalisation of conditioned fear in humans. *Journal of Behavior Therapy and Experimental Psychiatry*, *46*, 59–65. http://doi.org/10.1016/j.jbtep.2014.08.001

Arch, J. J., & Craske, M. G. (2008). Acceptance and commitment therapy and cognitive behavioral therapy for anxiety disorders: Different treatments, similar mechanisms? *Clinical Psychology: Science and …*, 263–280. Retrieved from http://onlinelibrary.wiley.com/doi/10.1111/j.1468-2850.2008.00137.x/full

Arch, J. J., Wolitzky-Taylor, K. B., Eifert, G. H., & Craske, M. G. (2012). Longitudinal treatment mediation of traditional cognitive behavioral therapy and acceptance and commitment therapy for anxiety disorders. *Behaviour Research and Therapy*, *50*(7-8), 469–78. http://doi.org/10.1016/j.brat.2012.04.007

Asparouhov, T., Muthén, L. K., & Muthén, B. O. (2014). Continuous - Time Survival Analysis in Mplus.

Augustson, E. M., & Dougher, M. J. (1997). The transfer of avoidance evoking functions through stimulus equivalence classes. *Journal of Behavior Therapy and Experimental Psychiatry*, *28*(3), 181–91. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/9327297

Bach, P., Gaudiano, B. a, Hayes, S. C., & Herbert, J. D. (2012). Acceptance and commitment therapy for psychosis: intent to treat, hospitalization outcome and mediation by believability. *Psychosis*, (April), 1–9. http://doi.org/10.1080/17522439.2012.671349

Barnes-Holmes, D., Hayes, S. C., Dymond, S., & O'Hora, D. (2001). Multiple stimulus relations and the transformation of stimulus functions. In S. C. Hayes, D. Barnes-Holmes, & B. Roche (Eds.), *Relational frame theory: A post-Skinnerian account of human language and cognition.* (pp. 51–71). New York, NY, US: Kluwer Academic/Plenum Publishers. Retrieved from http://ezproxy.gsu.edu/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=psyh&AN=2001-06976-003&site=ehost-live&scope=site

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*(2), 249–54. http://doi.org/10.1017/S1352465814000630

Berens, N. M., & Hayes, S. C. (2007). Arbitrarily applicable comparative relations: Experimental evidence for a relational operant. *Journal of Applied Behavior Analysis*, *40*(1), 45–71. Retrieved from http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=1868810&tool=pmcentrez&rendertype=abstract

Berkman, E. T., Lieberman, M. D., & Gable, S. L. (2009). BIS, BAS, and response conflict: Testing predictions of the revised reinforcement sensitivity theory. *Personality and Individual Differences*, *46*(5-6), 586–591. http://doi.org/10.1016/j.paid.2008.12.015

Berkowitz, L., & Donnerstein, E. (1982). External validity is more than skin deep: Some answers to criticisms of laboratory experiments. *American Psychologist*, *37*(3), 245–257. http://doi.org/10.1037/0003-066X.37.3.245

Biglan, A., & Hayes, S. C. (1997). Should the behavioral sciences become more pragmatic? The case for functional contextualism in research on human behavior. *Applied and Preventive Psychology*, *57*, 47–57. Retrieved from http://www.sciencedirect.com/science/article/pii/S0962184996800266

Blackledge, J. T. (2007). Disrupting verbal processes: Cognitive defusion in acceptance and commitment therapy and other mindfulness-based pscyhotherapies. *The Psychological Record*, *57*, 555–576. Retrieved from http://opensiuc.lib.siu.edu/tpr/vol57/iss4/6/

Blackledge, J. T., Moran, D. J., & Ellis, A. (2008). 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*(4), 232–248. http://doi.org/10.1007/s10942-007-0078-x

Bouton, M. E. (2002). Context, ambiguity, and unlearning: sources of relapse after behavioral extinction. *Biological Psychiatry*, *52*(10), 976–986. http://doi.org/10.1016/S0006-3223(02)01546-9

Bouton, M. E. (2004). Context and behavioral processes in extinction. *Learning & Memory (Cold Spring Harbor, N.Y.)*, *11*(5), 485–94. http://doi.org/10.1101/lm.78804

Bradley, M. M., Lang, P. J., Margaret, M., & Peter, J. (2007). The International Affective Digitized Sounds Affective Ratings of Sounds and Instruction Manual.

Carver, C., & White, T. (1994). Behavioral inhibition, behavioral activation, and affective responses to impending reward and punishment: The BIS/BAS Scales. *Journal of Personality and Social Psychology*, *67*(2), 319–313.

Chaplin, W. F. (1991). The next generation of moderator research in personality psychology. *Journal of Personality*, *59*(2), 143–78. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/1880698

Chawla, N., & Ostafin, B. (2007). Experiential avoidance as a functional dimensional approach to psychopathology: an empirical review. *Journal of Clinical Psychology*, *63*(9), 871–90. http://doi.org/10.1002/jclp.20400

Clark, L. A., & Watson, D. (1995). Constructing validity: Basic issues in objective scale development. *Psychological Assessment*, *7*(3), 309–319. http://doi.org/10.1037/1040-3590.7.3.309

Cortina, J. M. (1993). What is coefficient alpha? An examination of theory and applications. *Journal of Applied Psychology*, *78*(1), 98–104. http://doi.org/10.1037/0021-9010.78.1.98

Coxe, S., West, S. G., & Aiken, L. S. (2009). The analysis of count data: a gentle introduction to poisson regression and its alternatives. *Journal of Personality Assessment*, *91*(2), 121–36. http://doi.org/10.1080/00223890802634175

David, D., & Montgomery, G. H. (2011). The scientific status of psychotherapies: A new evaluative framework for evidencebased psychosocial interventions. *Clinical Psychology: Science and Practice*, *41*, 426–432.

Deacon, B. J., Fawzy, T. I., Lickel, J. J., & Wolitzky-Taylor, K. B. (2011). Cognitive Defusion Versus Cognitive Restructuring in the Treatment of Negative Self-Referential Thoughts: An Investigation of Process and Outcome. *Journal of Cognitive Psychotherapy*, *25*(3), 218–232. http://doi.org/10.1891/0889-8391.25.3.218

Dobson, K., & Beshai, S. (2013). The theory-practice gap in cognitive behavioral therapy: reflections and a modest proposal to bridge the gap. *Behavior Therapy*, *44*(4), 559–67. http://doi.org/10.1016/j.beth.2013.03.002

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*(2), 179–97. Retrieved from http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=1986433&tool=pmcentrez&rendertype=abstract

Dymond, S., Dunsmoor, J. E., Vervliet, B., & Roche, B. (2015). Fear Generalization in Humans: Systematic Review and Implications for Anxiety Disorder Research. *Behavior Therapy*, *46*(5), 561–582. http://doi.org/10.1016/j.beth.2014.10.001

Dymond, S., May, R. J., Munnelly, A., & Hoon, A. E. (2010). Evaluating the Evidence Base for Relational Frame Theory: A Citation Analysis. *The Behavior Analyst*, *33*(1), 97–117. Retrieved from http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2867509/

Dymond, S., Ng, T. C., & Whelan, R. (2013). Establishing Arbitrarily Applicable Relations of Same and Opposite With the Relational Completion Procedure: Selection-Based Feedback. *The Psychological Record*, *63*(1), 111–130. http://doi.org/10.11133/j.tpr.2013.63.1.009

Dymond, S., & Roche, B. T. (2009). A contemporary behavior analysis of anxiety and avoidance. *The Behavior Analyst*, (404). Retrieved from http://www.ncbi.nlm.nih.gov/pmc/articles/pmc2686994/

Dymond, S., Roche, B. T., Forsyth, J. P., Whelan, R., & Rhoden, J. (2007). Transformation of avoidance response functions in accordance with same and opposite relational frames. *Journal of the Experimental Analysis of Behavior*, *88*(2), 249–262. http://doi.org/10.1901/jeab.2007.22-07

Dymond, S., Roche, B. T., Forsyth, J. P., Whelan, R., & Rhoden, J. (2008). Derived avoidance learning: Transformation of avoidance response functions in accordance with same and opposite relational frames. *The Psychological Record*, *58*(2), 269–286. Retrieved from http://ezproxy.gsu.edu/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=psyh&AN=2008-05378-007&site=ehost-live&scope=site

Dymond, S., Schlund, M. W., Roche, B. T., De Houwer, J., & Freegard, G. P. (2012). Safe from harm: learned, instructed, and symbolic generalization pathways of human threat-avoidance. *PloS One*, *7*(10), e47539. http://doi.org/10.1371/journal.pone.0047539

Dymond, S., Schlund, M. W., Roche, B. T., & Whelan, R. (2014). The spread of fear: symbolic generalization mediates graded threat-avoidance in specific phobia. *Quarterly Journal of Experimental Psychology (2006)*, *67*(2), 247–59. http://doi.org/10.1080/17470218.2013.800124

Dymond, S., Schlund, M. W., Roche, B. T., Whelan, R., Richards, J., & Davies, C. (2011). Inferred threat and safety: symbolic generalization of human avoidance learning. *Behaviour Research and Therapy*, *49*(10), 614–21. http://doi.org/10.1016/j.brat.2011.06.007

Forman, E. M., Chapman, J. E., Herbert, J. D., Goetter, E. M., Yuen, E. K., & Moitra, E. (2012). Using session-by-session measurement to compare mechanisms of action for acceptance and commitment therapy and cognitive therapy. *Behavior Therapy*, *43*(2), 341–54. http://doi.org/10.1016/j.beth.2011.07.004

Forman, E. M., Herbert, J. D., Juarascio, A. S., Yeomans, P. D., Zebell, J. a., Goetter, E. M., & Moitra, E. (2012). The Drexel defusion scale: A new measure of experiential distancing. *Journal of Contextual Behavioral Science*, *1*(1-2), 55–65. http://doi.org/10.1016/j.jcbs.2012.09.001

Forsyth, J. P., Eifert, G. H., & Barrios, V. (2006). Fear conditioning in an emotion regulation context: A fresh perspective on the origins of anxiety disorders. In M. G. Craske, D. Hermans, & D. Vansteenwegen (Eds.), *Fear and learning: From basic processes to clinical applications* (pp. 133–153). Washington D.C.: American Psychological Association.

Gamez, W., Kotov, R., & Watson, D. (2010). The validity of self-report assessment of avoidance and distress. *Anxiety, Stress, and Coping*, *23*(1), 87–99. http://doi.org/10.1080/10615800802699198

Gavin, A., & Roche, B. T. (2008). Competing Contingencies Over Derived Relational Responding : A Behavioral Model of the Implicit Association Test, 427–441.

Gelfand, L. A., MacKinnon, D. P., DeRubeis, R. J., & Baraldi, A. N. (2016). Mediation Analysis with Survival Outcomes: Accelerated Failure Time vs. Proportional Hazards Models. *Frontiers in Psychology*, *7*, 423. http://doi.org/10.3389/fpsyg.2016.00423

Gillanders, D. T., Bolderston, H., Bond, F. W., Dempster, M., Flaxman, P. E., Campbell, L., … Remington, B. (2014). The Development and Initial Validation of the Cognitive Fusion Questionnaire. *Behavior Therapy*, *45*(1), 83–101. http://doi.org/10.1016/j.beth.2013.09.001

Goldin, P. R., Ziv, M., Jazaieri, H., Werner, K., Kraemer, H., Heimberg, R. G., & Gross, J. J. (2012). Cognitive reappraisal self-efficacy mediates the effects of individual cognitive-behavioral therapy for social anxiety disorder. *Journal of Consulting and Clinical Psychology*, *80*(6), 1034–40. http://doi.org/10.1037/a0028555

Gray, J. A. (1970). The psychophysiological basis of introversion-extraversion. *Behaviour Research and Therapy*, *8*(3), 249–266. http://doi.org/10.1016/0005-7967(70)90069-0

Gray, J. D., Hanna, D., Gillen, A., & Rushe, T. (2016). A closer look at Carver and White's BIS/BAS scales: Factor analysis and age group differences. *Personality and Individual Differences*, *95*, 20–24. http://doi.org/10.1016/j.paid.2016.02.022

Gross, J. J. (1998). Antecedent- and response-focused emotion regulation: divergent consequences for experience, expression, and physiology. *Journal of Personality and Social Psychology*, *74*(1), 224–37. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/9457784

Gross, J. J. (2002). Emotion regulation: Affective, cognitive, and social consequences. *Psychophysiology*, *39*. http://doi.org/10.1017/S0048577201393198

Hancock, G. R., & Mueller, R. O. (2010). Structural equation modelling. In G. R. Hancock & R. O. Mueller (Eds.), *The reviewer's guide to quantitative methods in the social sciences*. New York: Routledge.

Harmon, K., Strong, R., & Pasnak, R. (1982). Relational responses in tests of transposition with rhesus monkeys. *Learning and Motivation*, *13*(4), 495–504. Retrieved from http://ezproxy.gsu.edu/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=psyh&AN=1983-25001-001&site=ehost-live&scope=site

Hayes, A. F. (2013). *Introduction to mediation, moderation, and conditional process analysis: A regression-based approach*. New York: Guilford Press.

Hayes, A. F., & Preacher, K. J. (2014). Statistical mediation analysis with a multicategorical independent variable. *British Journal of Mathematical and Statistical Psychology*, *67*(3), 451–470. http://doi.org/10.1111/bmsp.12028

Hayes, S. C. (1976). The role of approach contingencies in phobic behavior. *Behavior Therapy*, *7*(1), 28–36. http://doi.org/10.1016/S0005-7894(76)80216-X

Hayes, S. C. (1994). Relational frame theory: A functional approach to verbal events. In S. C. Hayes, L. J. Hayes, M. Sato, & K. Ono (Eds.), *Behavior analysis of language and cognition.* (pp. 9–30). Reno, NV, US: Context Press. Retrieved from http://ezproxy.gsu.edu/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=psyh&AN=1994-98564-001&site=ehost-live&scope=site

Hayes, S. C. (2004). Acceptance and commitment therapy, relational frame theory, and the third wave of behavioral and cognitive therapies. *Behavior Therapy*, *35*(4), 639–665. http://doi.org/10.1016/S0005-7894(04)80013-3

Hayes, S. C. (2005). Stability and Change in Cognitive Behavior Therapy: Considering the Implications of ACT and RFT. *Journal of Rational-Emotive & Cognitive-Behavior Therapy*, *23*(2), 131–151. http://doi.org/10.1007/s10942-005-0007-9

Hayes, S. C. (2014). Re: [act] RE: [act] The classics [Electronic mailling list message]. Retrieved from https://contextualscience.org/emailing_lists

Hayes, S. C., Barnes-Holmes, D., & Roche, B. (Eds.). (2001). *Relational Frame Theory: A Post-Skinnerian Account of Human Language and Cognition*. New York, NY: Kluwer Academic Publishers. Retrieved from http://www.sciencedirect.com/science/article/pii/S0065240702800635/pdf?md5=31554347b560dd545331c5b86290cdc6&pid=1-s2.0-S0065240702800635-main.pdf&_valck=1

Hayes, S. C., Barnes-Holmes, D., & Wilson, K. G. (2012). Contextual Behavioral Science: Creating a science more adequate to the challenge of the human condition. *Journal of Contextual Behavioral Science*, *1*(1-2), 1–16. http://doi.org/10.1016/j.jcbs.2012.09.004

Hayes, S. C., & Berens, N. M. (2004). Why Relational Frame Theory Alters the Relationship between Basic and Applied Behavioral Psychology. *International Journal of Psychology and Psychological Therapy*, *4*, 341–353. Retrieved from http://www.ijpsy.com/volumen4/num2/88.html?accion=cambiar_idioma&lang=en\npapers2://publication/uuid/B7F71704-9C5B-4EFE-A0B8-B6DE66F455F5

Hayes, S. C., Bisset, R., Roget, N., Padilla, M., Kohlenberg, B. S., Fisher, G., … Niccolls, R. (2004). The Impact of Acceptance and Commitment Training and Multicultural Training on the Stigmatizing Attitudes and Professional Burnout of Substance Abuse Counselors. *Behavior Therapy*, *35*, 821–835.

Hayes, S. C., Blackledge, J. T., & Barnes-Holmes, D. (2001). Language and cognition: Constructing an alternative approach within the behavioral tradition. In S. C. Hayes, D. Barnes-Holmes, & B. Roche (Eds.), *Relational frame theory: A post-Skinnerian account of human language and cognition.* (pp. 3–20). New York, NY, US: Kluwer Academic/Plenum Publishers. Retrieved from http://ezproxy.gsu.edu/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=psyh&AN=2001-06976-001&site=ehost-live&scope=site

Hayes, S. C., Fox, E. J., Gifford, E. V, Wilson, K. G., Barnes-Holmes, D., & Healy, O. (2001). Derived relational responding as learned behavior. In S. C. Hayes, D. Barnes-Holmes, & B. Roche (Eds.), *Relational frame theory: A post-Skinnerian account of human language and cognition.* (pp. 21–49). New York, NY, US: Kluwer Academic/Plenum Publishers. Retrieved from http://ezproxy.gsu.edu/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=psyh&AN=2001-06976-002&site=ehost-live&scope=site

Hayes, S. C., Gifford, E. V, Townsend, R. C. J., & Barnes-Holmes, D. (2001). Thinking, problem-solving, and pragmatic verbal analysis. In S. C. Hayes, D. Barnes-Holmes, & B. Roche (Eds.), *Relational frame theory: A post-Skinnerian account of human language and cognition.* (pp. 87–101). New York, NY, US: Kluwer Academic/Plenum Publishers. Retrieved from http://ezproxy.gsu.edu/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=psyh&AN=2001-06976-005&site=ehost-live&scope=site

Hayes, S. C., Lattal, K. A., & Myerson, W. A. (1979). Strength of experimentally induced phobic behavior in rats: Avoidance versus dual-component formulations. *Psychological Reports*, *44*(3, Pt 1), 891–894. Retrieved from http://ezproxy.gsu.edu/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=psyh&AN=1980-29636-001&site=ehost-live&scope=site

Hayes, S. C., Levin, M., & Plumb-Vilardaga, 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*(2), 180–198. http://doi.org/10.1016/j.beth.2009.08.002.Acceptance

Hayes, S. C., Luoma, J. B., Bond, F. W., Masuda, A., & Lillis, J. (2006). Acceptance and commitment therapy: model, processes and outcomes. *Behaviour Research and Therapy*, *44*(1), 1–25. http://doi.org/10.1016/j.brat.2005.06.006

Hayes, S. C., & Smith, S. (2005). *Get Out of Your Mind and Into Your Life: The New Acceptance and Commitment Therapy*. Oakland, CA: New Harbinger Publications, Inc.

Hayes, S. C., Strosahl, K. D., & Wilson, K. G. (1999). *Acceptance and commitment therapy: An experiential approach to behavior change.* New York, NY, US: Guilford Press. Retrieved from http://ezproxy.gsu.edu/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=psyh&AN=1999-04037-000&site=ehost-live&scope=site

Hayes, S. C., Strosahl, K. D., & Wilson, K. G. (2012). *Acceptance and commitment therapy: The process and practice of mindful change* (2nd ed.).

Hayes, S. C., Villatte, M., Levin, M., & Hildebrandt, M. (2011). Open, aware, and active: contextual approaches as an emerging trend in the behavioral and cognitive therapies. *Annual Review of Clinical Psychology*, *7*, 141–168. http://doi.org/10.1146/annurev-clinpsy-032210-104449

Hayes, S. C., & Wilson, K. G. (1993). Some applied implications of a contemporary behavior-analytic account of verbal events. *The Behavior Analyst*, *16*(2), 283–301. Retrieved from http://ezproxy.gsu.edu/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=psyh&AN=1994-23852-001&site=ehost-live&scope=site

Healy, H., Barnes-holmes, Y., Barnes-Holmes, D., Keogh, C., Luciano, C., & Wilson, K. G. (2008). An experimental test of a cognitive defusion exercise: Coping with negative and positive self-statements, 623–640.

Heinze, G., & Schemper, M. (2002). A solution to the problem of separation in logistic regression. *Statistics in Medicine*, *21*(16), 2409–19. http://doi.org/10.1002/sim.1047

Herbert, J. D., & Forman, E. M. (2009). The Evolution of Cognitive Behavior Therapy, (1849), 3–25.

Herbert, J. D., & Forman, E. M. (2013). Caution: the differences between CT and ACT may be larger (and smaller) than they appear. *Behavior Therapy*, *44*(2), 218–23. http://doi.org/10.1016/j.beth.2009.09.005

Herbert, J. D., Gaudiano, B. a, & Forman, E. M. (2013). The importance of theory in cognitive behavior therapy: a perspective of contextual behavioral science. *Behavior Therapy*, *44*(4), 580–91. http://doi.org/10.1016/j.beth.2013.03.001

Hesser, H., Westin, V., Hayes, S. C., & Andersson, G. (2009). Clients' in-session acceptance and cognitive defusion behaviors in acceptance-based treatment of tinnitus distress. *Behaviour Research and Therapy*, *47*(6), 523–528. http://doi.org/10.1016/j.brat.2009.02.002

Hinton, M. J., & Gaynor, S. T. (2008). Cognitive defusion for psychological distress, dysphoria, and low self-esteem: A randomized technique evaluation trial of vocalizing strategies, *6*(3), 164–186.

Hofmann, S. G., Asmundson, G. J. G., & Beck, A. T. (2013). The science of cognitive therapy. *Behavior Therapy*, *44*(2), 199–212. http://doi.org/10.1016/j.beth.2009.01.007

Hofmann, S. G., Heering, S., Sawyer, A. T., & Asnaani, A. (2009). How to handle anxiety: The effects of reappraisal, acceptance, and suppression strategies on anxious arousal. *Behaviour Research and Therapy*, *47*(5), 389–94. http://doi.org/10.1016/j.brat.2009.02.010

Hooper, N., & McHugh, L. (2013). Cognitive defusion versus thought distraction in the mitigation of learned helplessness. *The Psychological Record*, *63*(1), 209–218. http://doi.org/10.11133/j.tpr.2013.63.1.016

Hooper, N., Saunders, J., & McHugh, L. (2010). The derived generalization of thought suppression. *Learning & Behavior*, *38*(2), 160–8. http://doi.org/10.3758/LB.38.2.160

Johnco, C., Wuthrich, V. M., & Rapee, R. M. (2014). The influence of cognitive flexibility on treatment outcome and cognitive restructuring skill acquisition during cognitive behavioural treatment for anxiety and depression in older adults: Results of a pilot study. *Behaviour Research and Therapy*, *57*, 55–64. http://doi.org/10.1016/j.brat.2014.04.005

Johnson, S. L., Turner, R. J., & Iwata, N. (2003). BIS/BAS levels and psychiatric disorder: An epidemiological study. *Journal of Psychopathology and Behavioral Assessment*, *25*(1).

Kanter, J. W. (2013). The vision of a progressive clinical science to guide clinical practice. *Behavior Therapy*, *44*(2), 228–33. http://doi.org/10.1016/j.beth.2010.07.006

Kashdan, T. B., Barrios, V., Forsyth, J. P., & Steger, M. F. (2006). Experiential avoidance as a generalized psychological vulnerability: Comparisons with coping and emotion regulation strategies. *Behaviour Research and Therapy*, *44*(9), 1301–1320. http://doi.org/10.1016/j.brat.2005.10.003

Kazdin, A. E. (1978). Evaluating the generality of findings in analogue therapy research. *Journal of Consulting and Clinical Psychology*, *46*, 673–686.

Keng, S.-L., Robins, C. J., Smoski, M. J., Dagenbach, J., & Leary, M. R. (2013). Reappraisal and mindfulness: a comparison of subjective effects and cognitive costs. *Behaviour Research and Therapy*, *51*(12), 899–904. http://doi.org/10.1016/j.brat.2013.10.006

Kishita, N., Muto, T., Ohtsuki, T., & Barnes-Holmes, D. (2014). Measuring the effect of cognitive defusion using the Implicit Relational Assessment Procedure: An experimental analysis with a highly socially anxious sample. *Journal of Contextual Behavioral Science, 3*(1), 8–15. http://doi.org/10.1016/j.jcbs.2013.12.001

Kleinbaum, D. G., & Klein, M. (2012). *Survival analysis: A self-learning text* (3rd ed.). New York: Springer.

Kraemer, H. C., Kiernan, M., Essex, M., & Kupfer, D. J. (2008). How and why criteria defining moderators and mediators differ between the Baron & Kenny and MacArthur approaches. *Health Psychology : Official Journal of the Division of Health Psychology, American Psychological Association*, *27*(2 Suppl), S101–8. http://doi.org/10.1037/0278-6133.27.2(Suppl.).S101

Lang, P. J., Bradley, M. M., & Cuthbert, B. N. (2008). International Affective Picture System (IAPS): Affective ratings of pictures and instruction manual.

Levin, M. E., Hildebrandt, M. J., Lillis, J., & Hayes, S. C. (2012). The impact of treatment components suggested by the psychological flexibility model: a meta-analysis of laboratory-based component studies. *Behavior Therapy*, *43*(4), 741–56. http://doi.org/10.1016/j.beth.2012.05.003

Long, J. S. (1997). *Regression models for categorical and limited dependent variables*. Sage Publications.

Lovibond, P. (2006). Fear and Avoidance: An Integrated Expectancy Model. In M. G. Craske, D. Hermans, & D. Vansteenwegen (Eds.), *Fear and learning: From basic processes to clinical implications.* (pp. 117–132). Washington, DC, US: American Psychological Association. Retrieved from http://ezproxy.gsu.edu/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=psyh&AN=2006-09216-006&site=ehost-live&scope=site

Luciano, C., Valdivia-Salas, S., Ruiz, F. J., Rodríguez-Valverde, M., Barnes-Holmes, D., Dougher, M. J., … Gutierrez, O. (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*(3-4), 120–134. http://doi.org/10.1016/j.jcbs.2013.05.001

Luciano, C., Valdivia-Salas, S., Ruiz, F. J., Rodríguez-Valverde, M., Barnes-Holmes, D., Dougher, M. J., … 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*(1), 94–111. http://doi.org/10.1002/jeab.68

MacKinnon, D. P., Fairchild, A. J., & Fritz, M. S. (2007). Mediation analysis. *Annual Review of Psychology*, *58*, 593–614. http://doi.org/10.1146/annurev.psych.58.110405.085542

Mandavia, A., & Masuda, A. (n.d.). A Cognitive Defusion Technique Decreases the Discomfort, Believability, and Decentering of Self- Referential Negative Body Image Thoughts.

Massaro, D. W., Venezky, R. L., & Taylor, G. A. (1979). Orthographic regularity, positional frequency, and visual processing of letter strings: a reply to Krueger's comments. *Journal of Experimental Psychology. General*, *108*(1), 131–132.

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*(4), 477–85. http://doi.org/10.1016/j.brat.2003.10.008

Masuda, A., Hayes, S. C., & Twohig, M. P. (2009). A Parametric Study of Cognitive Defusion and the Believability and Discomfort of Negative Self-Relevant Thoughts, 250–262.

Masuda, A., Twohig, M. P., Stormo, A. R., Feinstein, A. B., Chou, Y.-Y., & Wendell, J. W. (2010). The effects of cognitive defusion and thought distraction on emotional discomfort and believability of negative self-referential thoughts. *Journal of Behavior Therapy and Experimental Psychiatry*, *41*(1), 11–7. http://doi.org/10.1016/j.jbtep.2009.08.006

Mennin, D. S., Ellard, K. K., Fresco, D. M., & Gross, J. J. (2013). United we stand: emphasizing commonalities across cognitive-behavioral therapies. *Behavior Therapy*, *44*(2), 234–48. http://doi.org/10.1016/j.beth.2013.02.004

Mook, D. G. (1983). In defense of External Invalidity. *American Psychologist*, *38*(4), 379–387.

Muthén, L. K., & Muthén, B. O. (2012). *Mplus User's Guide* (7th ed.). Los Angeles: Muthén & Muthén.

Öst, L.-G. (2014). The efficacy of Acceptance and Commitment Therapy: An updated systematic review and meta-analysis. *Behaviour Research and Therapy*, *61*, 105–121. http://doi.org/10.1016/j.brat.2014.07.018

Pilecki, B. C., & Mckay, D. (2012). An experimental investigation of cognitive defusion. *The Psychological Record*, *62*, 19–40.

Roche, B. T., Dymond, S., Melia, R., Kanter, J. W., & Blackledge, J. T. (2009). *The effect of a laboratory analog defusion procedure on the derived transfer of avoidance functions*.

Roche, B. T., Kanter, J. W., Brown, K. R., Dymond, S., & Fogarty, C. C. (2008). A comparison of "direct" versus "derived" extinction of avoidance responding. *The Psychological Record*, *58*(3), 443–464. Retrieved from http://ezproxy.gsu.edu/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=psyh&AN=2008-11965-007&site=ehost-live&scope=site

Selig, J. P., Preacher, K. J., & Little, T. D. (2012). Modeling Time-Dependent Association in Longitudinal Data: A Lag as Moderator Approach. *Multivariate Behavioral Research*, *47*, 697–716. http://doi.org/10.1080/00273171.2012.715557

Sidman, M. (1971). Reading and auditory-visual equivalences. *Journal of Speech & Hearing Research*, *14*(1), 5–13. Retrieved from http://ezproxy.gsu.edu/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=psyh&AN=1972-21539-001&site=ehost-live&scope=site

Sidman, M., & Tailby, W. (1982). Conditional discrimination vs. matching to sample: An expansion of the testing paradigm. *Journal of the Experimental Analysis of Behavior*, *37*(1), 5–22. Retrieved from http://ezproxy.gsu.edu/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=psyh&AN=1982-25577-001&site=ehost-live&scope=site

Skinner, B. F. (1957). *Verbal behavior.* East Norwalk, CT, US: Appleton-Century-Crofts. Retrieved from http://ezproxy.gsu.edu/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=psyh&AN=2006-20887-000&site=ehost-live&scope=site

Snyder, K., Lambert, J., & Twohig, M. P. (2011). Defusion: a behavior-analytic strategy for addressing private events. *Behavior Analysis in Practice*, *4*(2), 4–13. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/22649574

Stewart, I., Barnes-Holmes, D., Hayes, S. C., & Lipkens, R. (2001). Relations among relations: Analogies, metaphors, and stories. In S. C. Hayes, D. Barnes-Holmes, & B. Roche (Eds.), *Relational frame theory: A post-Skinnerian account of human language and cognition.* (pp. 73–86). New York, NY, US: Kluwer Academic/Plenum Publishers. Retrieved from http://ezproxy.gsu.edu/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=psyh&AN=2001-06976-004&site=ehost-live&scope=site

Swain, J., Hancock, K., Hainsworth, C., & Bowman, J. (2013). Acceptance and Commitment Therapy in the treatment of anxiety: A systematic review. *Clinical Psychology Review*, *33*(8), 965–978. http://doi.org/10.1016/j.cpr.2013.07.002

Törneke, N. (2010). *Learning RFT: An introduction to relational frame theory and its clinical applications*. Oakland, CA: Context Press/New Harbinger Publications. Retrieved from http://ezproxy.gsu.edu/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=cat00477a&AN=gast.2243128&site=eds-live

Troy, A. S., Wilhelm, F. H., Shallcross, Amanda, J., & Mauss, I. B. (2010). Seeing the Silver Lining: Cognitive Reappraisal Ability Moderates the Relationship Between Stress and Depressive Symptoms. *Emotion*, *10*(6), 783–795. http://doi.org/10.1037/a0020262.Seeing

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*(1), 85–111. http://doi.org/10.1901/jeab.2009.92-85

VanderWeele, T. J. (2011). Causal Mediation Analysis With Survival Data. *Epidemiology*, *22*(4), 582–585. http://doi.org/10.1097/EDE.0b013e31821db37e

Vilardaga, R., Hayes, S. C., Levin, M. E., & Muto, T. (2009). Creating a strategy for progress: a contextual behavioral science approach. *The Behavior Analyst / MABA*, *32*(1), 105–33. Retrieved from http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2686981&tool=pmcentrez&rendertype=abstract

Villatte, M. (2014). Re: [act] RE: [act] The classics [Electronic mailling list message].

Wegner, D. M. (1994). Ironic processes of mental control. *Psychological Review*, *101*(1), 34–52. http://doi.org/10.1037/0033-295X.101.1.34

Whelan, R., Cullinan, V., Donovan, A. O., & Valverde, M. R. (2005). Derived Same and Opposite Relations Produce Association and Mediated Priming, *5*, 247–264.

Wilamowska, Z. A., Thompson-Hollands, J., Fairholme, C. P., Ellard, K. K., Farchione, T. J., & Barlow, D. H. (2010). Conceptual background, development, and preliminary data from the unified protocol for transdiagnostic treatment of emotional disorders. *Depression and Anxiety*, *27*(10), 882–90. http://doi.org/10.1002/da.20735

Wilson, K. G., & Hayes, S. C. (1996). Resurgence of derived stimulus relations. *Journal of the Experimental Analysis of Behavior*, *66*(3), 267–81. http://doi.org/10.1901/jeab.1996.66-267

Wilson, K. G., Hayes, S. C., Gregg, J., & Zettle, R. D. (2001). Psychopathology and psychotherapy. In S. C. Hayes, D. Barnes-Holmes, & B. Roche (Eds.), *Relational frame theory: A post-Skinnerian account of human language and cognition.* (pp. 211–237). New York, NY, US: Kluwer Academic/Plenum Publishers. Retrieved from http://ezproxy.gsu.edu/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=psyh&AN=2001-06976-012&site=ehost-live&scope=site

Wilson, K. G., & Murrell, A. R. (2004). Values work in acceptance and commitment therapy: Setting a course for behavioral treatment. In S. C. Hayes, V. C. Follette, & M. M. Linehan (Eds.), *Mindfulness and acceptance: Expanding the cognitive-behavioral tradition* (pp. 120–151). New York: Gilford.

Wolgast, Martin, Lundh, Lars-Gunnar, Viborg, G. (n.d.). Experiential Avoidance as an Emotion Regulatory Function: An Empirical Analysis of Experiential Avoidance in Relation to Behavioral Avoidance, Cognitive Reappraisal, and Response Suppression.

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

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

Wulfert, E., & Hayes, S. C. (1988). Transfer of a conditional ordering response through conditional equivalence classes. *Journal of the Experimental Analysis of Behavior, 50*(2), 125–144.

Yovel, I., Mor, N., & Shakarov, H. (2014). Examination of the core cognitive components of cognitive behavioral therapy and acceptance and commitment therapy: an analogue investigation. *Behavior Therapy*, *45*(4), 482–94. http://doi.org/10.1016/j.beth.2014.02.007

Zettle, R. D. (2011). The Evolution of a Contextual Approach to Therapy : From Comprehensive Distancing to ACT, *7*(1), 76–82.

Zettle, R. D., & Hayes, S. C. (1986). Dysfunctional control by client verbal behavior: The context of reason-giving. *The Analysis of Verbal Behavior*, *4*, 30–38. Retrieved from http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2748463&tool=pmcentrez&rendertype=abstract

Zettle, R. D., Rains, J. C., & Hayes, S. C. (2011). Processes of change in acceptance and commitment therapy and cognitive therapy for depression: a mediation reanalysis of Zettle and Rains. *Behavior Modification*, *35*(3), 265–83. http://doi.org/10.1177/0145445511398344

APPENDICES

Appendix A

Defusion Protocol

Rationale

Now I am going to train you on how to use a strategy for dealing with negative thoughts. Much of thinking is really self-talk using words. This is an incredibly useful ability but it can cause problems. For example, imagine you experience the thought “I am a failure”. It could be a passing thought, with very little basis in reality, but if you automatically buy into the thought, this collection of words “I - am - a – failure” could immediately make you feel sad our guilty. The good news is that we can choose to interact differently with our thoughts. Words don’t inherently have power to hurt us. In reality, they are just arbitrary collections of sounds or symbols. For example, the word failure is really just a collection of sounds: “fail-ure” (*said slowly).* It is our minds that actively give the word meaning and psychological impact. This usually happens so automatically that we don’t even notice the process. However, we can disrupt the process if we treat our thoughts as “just thoughts” – words, sounds and symbols, temporarily floating around in our minds. For example, imagine you have the thought, “I am a failure,” but this time you recognize it as just words floating in your mind, which you do not have to accept or believe. By holding the thought lightly, the impact may be lessened. In summary, we can lessen the impact of thoughts by understanding that they are simply collections of words, sounds, and images that may or may not have any basis in reality. I realize that what I’m describing may seem a bit abstract, so let’s try a little exercise that should make things more clear.

Training

I'm going to ask you to say a word. Then you tell me what comes to mind. I want you to say the word, "milk".

Brief pause

Good. Now tell me what comes to mind when you said it?

Brief pause

O.K. what else? What shows up when we say "milk"?

Brief pause

[if participant does not come up with the examples below]
Good what else?
Can you picture the color?
Can you taste it?

Can you feel what it feels like to drink a glass of milk?
Can you feel the temperature, the texture, etc.

Brief pause

O.K. let's see if this fits. What came across your mind were things about actual milk and your experience with it. All that happened is that we made a strange sound — Milk *(said slowly)* — and lots of those things show up. Notice that there isn't any milk in this room, not at all. But milk was in the room **psychologically**. You and I could see it, taste it, and feel it. And yet, only the word was actually here.

Brief pause

Now, here is another exercise. I am going to do it with you. What I am going to ask you to do is to say the word, "milk," out loud, over-and-over again, and as rapidly as possible, and then notice what happens to the image and feelings associated with milk. Are you ready? O.K., Let's do it. Say "milk" over and over again.

(*Experimenter repeats the word with the participant initially to prompt him or her to follow the protocol AND then provide a prompt every 10 seconds, by saying "louder" or "faster" etc*).

Repeat "milk" rapidly for 20 seconds

O.K., Now stop. Tell me what came to mind while you kept repeating it?

Brief pause

Did you notice what happened to the images and feelings associated with milk that were here a few minutes ago?

Brief pause

OK. When you said it the first time, it was as if milk was actually here, in the room. But all that really happened was that you merely said a word. The first time you said it, it was "*psychologically*" meaningful, and it was almost real. But when you said it again and again and again, you began to lose that meaning and the word became just a sound.

What I am suggesting is that what happened in this exercise may be applied to other words and thoughts that you contact. The words and thoughts are like the word "milk"…a collection of sounds or symbols that do not inherently have the power to impact you psychologically.

Exercise. Let's do another exercise with one of the words you saw in the experiment. We will use the word [*insert C1*]. When you contacted this word you likely experienced particular thoughts or emotions.

What are some thoughts and feelings that you have when you hear this word?

Just like we did with "milk," I would like you to repeat [*insert C1*] out loud, over and over again, as rapidly as possible until I say "stop".

O.K., are you ready? Now, begin.

(*Experimenter repeats the word with the participant initially to prompt him or her to follow the protocol AND then provide a prompt every 10 seconds, by saying "louder" or "faster" etc*).

Participant repeats word for 30 seconds

Okay, stop.
Did you notice a change in the images and feelings associated with [Insert C1]?

Brief pause

Notice if you began to lose the meaning and psychological impact your mind attached to the word [Insert C1], and like the word "milk" it became just a collection of sounds and symbols.

When you contact [Insert C1], notice if this exercise helps reduce the automatic psychological impact your mind attaches to [Insert C1] and see if you can experience [Insert C1] as just a collection of sounds and symbols.

Appendix B

Reappraisal Protocol

Rationale

Now I am going to train you on how to use a strategy for dealing with negative emotions. We call this technique "reappraisal". What this technique involves is to change our thoughts or interpretations about events that upset us—meaning—to change the way we think about the event. Typically, the reason we experience difficult thoughts and emotions is that we tend to have very negative interpretations about the event that triggers them. The technique of reappraisal is to change the way we interpret the event, so that we feel less negatively about it. It's the attempt to develop a more positive interpretation about a situation that, at first glance, may appear very negative to us. There are many ways we can change our interpretations about a situation. We can start, for example, by asking 'Is this the only way I can think about the event or the situation?' 'Is there a more positive way of construing the event so that I feel less negatively about it?,' or 'What could be another way to look at what happened?' If you like, you can also challenge your negative thoughts or interpretations of an event, disprove them by reasoning or by drawing from your other experiences, and replace them with more positive interpretations.

Training

Let's do a practice exercise to demonstrate how the technique of reappraisal works with difficult experiences. Let's say you have a colleague—a classmate, a coworker, or a friend—with whom you have a fairly good relationship. Each time you run into her, she always smiles or says hi. Imagine, though, that one day you run into her, and she isn't smiling at you. In fact, she may be even frowning or scowling and when you try to say hi, she looks away and turns the corner.

What might be some negative thoughts and feelings that you would have in that moment? Try to bring those thoughts and feelings to mind, to really experience them.

Now, to begin practicing reappraisal I'm going to ask you to bring to mind one negative thought that you might have in that situation—when your colleague isn't smiling at you like normal.

Brief pause

[If participant does not come up with a negative thought/interpretation]*: One negative interpretation of this situation would be, "My friend ignored me on purpose." If you interpreted the situation this way, you might even think that you did something to offend your friend or make them angry. You might feel angry or insulted, or perhaps sad or even lonely.*

Now, ask yourself how the event is being interpreted, and how could it be interpreted differently. One way to reinterpret this event is to think about the event from a perspective other than one's own, say an impartial observer or a third person. Try asking yourself, "If another person, who didn't know your friend or colleague, were to view the event unfold, how might they interpret the situation? What sort of advice might they give you?"

Pause

*[*if person does not offer up ideas] *A third person might notice your friend's behavior and think "Your friend must be having a bad day," or "Your friend was probably distracted or in a hurry," If you interpreted the event this way, you might think about what was happening in your friend's life. You might feel concerned for your friend, or you might simply not get as upset.*

As you can see, adopting a different interpretation of a negative situation might change how we feel about it.

To summarize, the purpose of this technique is to think about the situation that upsets us in a different way, or cast it in more positive light, so that we are less emotionally impacted by it.

Exercise. Let's do another reappraisal exercise with one of the words you saw in the last part of the experiment. We will use the word [*insert C1*]. When you contacted this word you likely experienced particular thoughts or emotions.

What are some thoughts and feelings that you have when you hear this word?

Brief pause.

I would like you to practice reappraisal with the word. I am going to hold up a card and say the word. As I do so try to adopt a detached and unemotional attitude. In other words, as I say the word, try to think about what you are seeing and hearing objectively, in terms of the technical aspects of the word. Similar to the practice exercise, you may choose to adopt the perspective of another person, such as someone who was not part of this experiment and had no experience with the word. How might they interpret the word? Attend to the word carefully, but please try to think about it in such a way that you don't feel anything at all.

Experimenter repeats the word

Notice if adopting a different perspective changed how you feel about the [*insert C1*].

When you contact this word again notice if adopting a different perspective or interpretation changes your feelings associated with it.

Appendix C

Control

I am going to ask you some questions about your daily life and would like you to come up with at least two examples per question.

What do you usually do on campus?

Pause for participant's response. Query if less than two examples given.

Do you have any hobbies?

Pause for participant's response. Query if less than two examples given.

What kind of books do you like to read?

Pause for participant's response. Query if less than two examples given.

What kind of TV shows do you like to watch?

Pause for participant's response. Query if less than two examples given.

What kind of movies do you like to watch?

Pause for participant's response. Query if less than two examples given.

What do you usually do during holiday vacations?

Pause for participant's response. Query if less than two examples given.

Where did you go to high school?

Pause for participant's response.

What did you like about your high school?

Pause for participant's response. Query if less than two examples given.

What would you change about your high school?

Pause for participant's response. Query if less than two examples given.

Do you have a major?

[if no] What major are you considering?

What do you like about that major?

Pause for participant's response. Query if less than two examples given.

What would you change about that major?

Pause for participant's response. Query if less than two examples given.

What do you like about this University?

Pause for participant's response. Query if less than two examples given.

What would you change about this University?

Pause for participant's response. Query if less than two examples given.

What is your living situation?

Pause for participant's response.

What do you like about your living situation?

Pause for participant's response. Query if less than two examples given.

What would you change about your living situation?

Pause for participant's response. Query if less than two examples given.

What do you like about this city?

Pause for participant's response. Query if less than two examples given.

What don't you like about this city?

Pause for participant's response. Query if less than two examples given.