8-1-2018

TASTING WHAT YOU SEE: USING THE IMPLICIT RELATIONAL ASSESSMENT PROCEDURE TO MEASURE THE EFFECTS OF THE WORD REPETITION TECHNIQUE

Travis Sain
Southern Illinois University Carbondale, travis.n.sain@gmail.com

TASTING WHAT YOU SEE: USING THE IMPLICIT RELATIONAL ASSESSMENT PROCEDURE TO MEASURE THE EFFECTS OF THE WORD REPETITION TECHNIQUE

by

Travis Sain

B.S., Grand Valley State University, 2011
M.A., Southern Illinois University Carbondale, 2015

A Dissertation
Submitted in Partial Fulfillment of the Requirements for the
Doctor of Philosophy degree

Department of Psychology
in the Graduate School
Southern Illinois University Carbondale
August 2018

DISSERTATION APPROVAL

TASTING WHAT YOU SEE: USING THE IMPLICIT RELATIONAL ASSESSMENT PROCEDURE TO MEASURE THE EFFECTS OF THE WORD REPETITION TECHNIQUE

by

Travis Sain

A Dissertation Submitted in Partial

Fulfillment of the Requirements

for the Degree of

Doctor of Philosophy

in the field of Clinical Psychology

Approved by:

Chad E. Drake, Chair

Sarah Kertz

Reza Habib

Holly Cormier

Ruth Anne Rehfeldt

Graduate School
Southern Illinois University Carbondale
May 21, 2018

AN ABSTRACT OF THE DISSERTATION OF

TRAVIS SAIN, for the Doctor of Philosophy degree in Clinical Psychology, presented on May 21, 2018, at Southern Illinois University Carbondale.

TITLE: TASTING WHAT YOU SEE: USING THE IMPLICIT RELATIONAL ASSESSMENT PROCEDURE TO MEASURE THE EFFECTS OF THE WORD REPETITION TECHNIQUE

MAJOR PROFESSOR: Chad E. Drake, Ph.D.

The current study tested the effects of a cognitive defusion intervention on implicit attitudes toward milk and lemon as measured by the Implicit Relational Assessment Procedure (IRAP). One-hundred and eleven participants were randomly assigned to one of four conditions: control math, control defusion, half defusion, or full defusion. Participants filled out a series of self-report measures at the beginning of the study on psychological functioning, as well as attitudes toward milk and lemon. Participants then completed a task specific to their condition, with control math participants completing a simple math task while defusion conditions completed a defusion intervention – word repetition technique (WRT) – for certain words. The control defusion condition completed the WRT for the words "car" and "rabbit," the half defusion condition completed the WRT for the word "milk," and the full defusion condition completed the WRT for the words "milk" and "lemon." After completing the condition specific tasks, all participants completed a milk/lemon IRAP that included the words "milk" and "lemon" and pictures of milk and lemon. All participants finished the study by completing a final set of self-report measures. Results of the study indicated that IRAP performance was not significantly different between conditions following various levels of a defusion intervention. However, results showed that the pattern of IRAP response latencies did significantly vary between conditions, but this effect was driven by a significant difference on a single response latency between two conditions suggesting this finding is an artifact. Thus, the current study cannot

conclude that a defusion intervention can significantly affect implicit attitudes towards common objects, and any future research should consider applying a defusion intervention to clinically relevant stimuli to further assess for defusion effects in the IRAP.

TABLE OF CONTENTS

LIST OF TABLES

CHAPTER 1

INTRODUCTION

Mental health is one of the most costly and widespread issues around the world, with estimated costs for treating mental health concerns in major industrialized nations reaching billions of dollars per year for individual countries (Schofield et al., 2011). Furthermore, these estimates do not take into account the "costs" of psychological disorders on other facets of individuals' lives, including workplace performance, physical health, utilization of emergency and medical services, substance abuse, and funeral costs, among others (Schofield et al., 2011). For example, employees suffering from various levels of psychological distress contribute to a loss of billions of dollars each year in the workplace, and these financial losses would be substantially larger if one considered the worldwide number of employees dealing with diagnosed and undiagnosed psychological disorders (Birnbaum et al., 2010). Estimates indicate that lifetime prevalence rates (LPR) for various psychological disorders vary widely among depressive and anxiety disorders, with a LPR for specific phobia at 15.6%, social phobia at 10.7%, posttraumatic stress disorder (PTSD) at 5.7%, depressive episode at 16.6%, and major depression at 14.4% (Kessler, Petukhova, Sampson, Zaslavsky, & Wittchen, 2012). These LPR estimates do not include former axis two diagnoses that contribute to high utilization of emergency and medical services (Maclean, Xu, French, & Ettner, 2014), with estimates of a single diagnosis of a personality disorders reaching as high as 23% for males and 20% for females in the National Epidemiological Survey on Alcohol and Related Conditions (Maclean et al., 2014). Thus, the evidence indicates that millions of individuals around the world are suffering with concerns regarding their mental health, and these concerns continue to contribute

to growing financial costs associated with the treatment of mental health problems, as well as other financial burdens related to these disorders.

While psychological disorders vary in the presence and presentation of symptoms, many psychological disorders have high comorbidity rates with one another. For example, Rodriguez et al. (2004) noted that in many clinical samples 50% or more of individuals experience an additional anxiety or mood disorder after receiving a primary diagnosis of an anxiety disorder, with comorbidity rates being even higher depending on the disorders being diagnosed. In their sample of primary care patients, Rodriguez et al. (2004) found that over 70% of primary care patients met diagnostic criteria for multiple disorders, and more than 90% of patients experienced multiple psychological disorders throughout their lifetime. Not surprisingly, mood and anxiety disorders occur in up to 78% of patients throughout their lives (Rodriguez et al., 2004). High comorbidity may present a challenge to the veracity of the diagnostic system or suggest a common factor across diagnoses. One possible common factor is the presence of dysfunctional or problematic cognitions that contribute to an individual's distress, a symptom that is listed for many mood, anxiety, and personality disorders.

For example, a common problematic thought in individuals suffering with depression may be some variation of the self-directed judgment "I am worthless." When an individual experiences this kind of thought, the response to that thought may be regarded as functional and helpful (e.g., engaging in pleasurable activities, interacting with people, exercising, etc.) or dysfunctional and unhelpful (e.g., isolating self from others, engaging in unhealthy habits such as poor appetite or sleep patterns, living a sedentary lifestyle, etc.). Engaging in productive behaviors in response to the thought "I am worthless" may act as a moderating effect and reduce the likelihood that depressive symptoms will continue to develop, while engagement in

unhealthy behaviors may reinforce this thought as well as increase the likelihood that other depressive symptoms will appear and be maintained. Similarly, individuals who experience anxiety may experience problematic thoughts like "if I feel anxious something is wrong with me" or "if I talk about my feelings others will reject me." Again, individuals who respond productively to such thoughts (e.g., being open to unpleasant feelings, engaging authentically with others, etc.) may be less likely to develop other symptoms of anxiety disorders. In contrast, responding dysfunctionally (e.g., suppressing unpleasant emotions, working excessively for approval from others, etc.) may negatively reinforce their anxiety and increase their chances of developing and maintaining an anxiety disorder. Therefore, based on these situations and countless other similar scenarios regarding problematic thoughts, how individuals respond to their problematic thoughts can determine whether or not one develops and/or maintains a diagnosable psychological disorder. The field of psychotherapy includes numerous evidence-based techniques and therapies to help individuals respond productively to problematic thoughts, or to help individuals experiencing clinically significant symptoms return to more productive behaviors in response to their cognitive experiences.

Empirically Supported Treatments for Problematic Cognitions

Over the last half century psychology has continued to develop and promote the use of empirically supported treatments for the full gamut of psychological disorders. Since the cognitive revolution in psychology during the 1950s and 1960s, researchers and clinicians alike have become increasingly well-versed in the influences of one's mind on their behavior (Leahey, 2001). Yet, professionals in the field have diverged on many fronts regarding the role of the mind in contributing to psychopathology, as well as how best to intervene in problematic cognitions. Recent decades have seen numerous therapies that take different approaches to the

treatment of problematic cognitions receive empirical support, despite their varied theoretical underpinnings regarding cognition. Below, a description of two of the most influential theories in the last 10 years are provided, as well as a brief review of the empirical research on the success of the active components of each treatment for problematic cognitions.

Cognitive Behavior Therapy. Since Beck (e.g., 1967) began publishing his cognitive theory of depression, which later expanded to Cognitive Behavior Therapy (CBT), CBT has become the most empirically supported treatment for a host of psychological disorders - especially disorders emphasizing problematic cognitions. Additionally, many variants of CBT have been developed over the decades, such as cognitive processing therapy and prolonged exposure for PTSD (Institute of Medicine, 2008), and the Unified Protocol for Transdiagnostic Treatment of Emotional Disorders (Ellard, Fairholme, Boisseua, Farchione, & Barlow, 2010). In general CBT favors well in comparison to other forms of therapy, surpassing treatment outcomes for pharmacotherapy, behavior therapy, and psychodynamic therapy in a meta-analysis for the treatment of depression, generalized anxiety, panic, and obsessive-compulsive disorder (Butler, Chapman, Forman, & Beck, 2006). In addition, Hofman, Asnaani, Vonk, Sawyer, and Fang (2012) conducted a meta-analysis of over 260 meta-analytic studies for CBT and found CBT to be effective in comparison to other treatments for substance use disorders, psychotic disorders, depression and mood disorders, anxiety disorders, bulimia, and some personality disorders. These studies and many others contribute to the growing sense that CBT is the first-line treatment for the vast majority of psychological disorders.

Beck (e.g., 1967) first began espousing the cognitive model in the 1960s as it relates to the treatment of depression, and the model has since evolved into modern conceptions of CBT. At its core, the cognitive model posits that thoughts, emotions, and behaviors are linked to one

another due to the effects of distorted cognitions on one's emotions and behaviors (Beck 1995; Beck 2011). In other words, a problematic thought precedes and is presumed to cause a subsequent negative emotion or behavioral response. For example, if an individual has the thought "Nobody loves me," it may be followed by distressing emotions such as sadness and anxiety, as well as dysfunctional behaviors like isolation from others or excessive efforts to manage social perceptions of self. Therefore, within this scenario a clinician may focus on intervening upon the problematic cognition(s), which have also been referred to as cognitive distortions or negative automatic thoughts throughout the years (Beck, 1967; Beck, 2011). To achieve this goal CBT clinicians utilize cognitive restructuring techniques to alter problematic cognitions, or as Beck (1967) stated, neutralize negative automatic thoughts. Beck (1967) elaborated on this concept by stating that if an individual can recognize their problematic thoughts and respond to them in an objective manner s/he will be able to mitigate the effects of such thoughts. As practiced today, cognitive restructuring has four steps: (1) identification of a problematic cognition/negative automatic thoughts, (2) recognition of the distorted thinking pattern leading to the negative automatic thought, (3) objectively and rationally challenge the negative automatic thought, and (4) generate alternative thoughts that incorporate the evidence created in step three (Hope, Burns, Hayes, Herbert, & Warner, 2010). Inherent within this approach, and the cognitive model of CBT more broadly, is the assumption that if one wants to overcome their problematic cognitions they must learn to use cognitive restructuring (Heimberg & Barlow, 1991). However, empirical research has not fully supported this assumption.

Empirical findings for cognitive restructuring. Through the use of various methodological comparisons (e.g., additive or dismantling approaches), findings on the effectiveness for cognitive restructuring have been inconsistent. Numerous studies have

demonstrated the additive effects of cognitive restructuring with other therapeutic techniques, such as studies showing the addition of cognitive restructuring with exposure therapy for patients with PTSD lead to better treatment outcomes than exposure therapy alone (Bryant, Moulds, Guthrie, Dang, & Nixon, 2003; Bryant et al., 2008). Another study conducted by Mattick and Peters (1988) tested whether the inclusion of cognitive restructuring with guided exposure in severe phobic patients would lead to better treatment outcomes. Results showed that participants who received cognitive restructuring and exposure produced significantly better end of treatment functioning than participants who received only exposure. In a comparison of behavior activation (BA), cognitive therapy (CT), and antidepressant medications (ADM) for depression, Dobson et al. (2008) found that CT evidenced stronger relapse prevention and recurrence of symptoms than BA and ADM, although differences between the BA and CT conditions were non-significant. Such studies demonstrate support for the claim that cognitive restructuring (or CT more broadly) may provide treatment effects that cannot be achieved by behavior therapy techniques alone, although many studies contest this finding.

In a study conducted by Biran and Wilson (1981), guided exposure was compared to cognitive restructuring for treating three different phobias (heights, elevators, and darkness). Biran and Wilson (1981) found that exposure demonstrated significantly better post-treatment self-efficacy, physiological responses, and approach behaviors to phobic scenarios, and increased ability to cope with phobic situations in day-to-day life. Hope, Heimberg, and Bruch (1995) compared a group of subjects receiving a group CBT treatment to subjects receiving exposure therapy and found that the group CBT was generally less effective than exposure alone for social phobia, and any treatment effects specific to group CBT were no longer present at 6-month follow-up. In a meta-analysis comparing cognitive-behavioral and pharmacological treatments

for social phobia, Gould, Buckminster, Pollack, Otto, and Yap (1997) found that exposure therapy alone produced the largest effect sizes (.89), and when cognitive restructuring is combined with exposure the treatment is still effective but demonstrates a smaller effect size (.80). Similarly, several studies have found that the addition of cognitive restructuring or CT to exposure therapy produced no additive effects in the treatment of social phobia, and that exposure therapy alone outperformed conditions including CT components (Salaberria & Echeburua, 1998; Scholing & Emmelkamp, 1996; Taylor et al., 1997). van Dam-Baggen and Kraaimaat (2000) conducted a study on individuals with generalized social phobia in which they compared social skills training (SST) to group CBT and found that SST outcomes exceeded those produced by group CBT. Additionally, Jacobson et al. (2000) conducted a dismantling study for treating depression with various components of CBT. Jacobson and colleagues found that BA alone, BA with training on managing automatic thoughts, and a full CBT treatment (with cognitive restructuring) produced equivalent outcomes immediately post-treatment, as well as at 6-month follow-up. Furthermore, Jacobson et al. (2000) noted that BA and training to manage automatic thoughts was as effective as CT for altering problematic thought patterns, suggesting that cognitive-specific treatments may not be necessary for treating problematic thoughts in CBT.

As the above research suggests, empirical support for the use of cognitive restructuring and cognitive therapy techniques has been inconsistent over the last few decades. Many studies have demonstrated that cognitive therapy is effective in the treatment of various psychological disorders (e.g., Bryant et al., 2003; 2008; Mattick & Peters, 1988), while numerous others lack supporting evidence on the efficacy of cognitive restructuring's additive influence in therapy (e.g., Biran & Wilson, 1981; Gould et al., 1997). Furthermore, multiple studies have shown that

behavior therapy techniques can produce treatment effects that were once believed to require cognitive restructuring to achieve (e.g., Jacobson et al., 2000; Longmore & Worrell, 2007), suggesting that cognitive restructuring may not be the most effective means – or at least an inconsistent way – of treating problematic cognitions across different presentations of psychopathology.

Acceptance and Commitment Therapy. An alternative to CBT, Acceptance and Commitment Therapy (ACT, said as a word: Hayes, Strosahl, & Wilson, 2012), has received empirical support for treatment of a wide variety of psychological concerns. ACT is a transdiagnostic approach that conceptualizes human suffering (i.e., psychological distress) as an excess of psychological inflexibility or a deficit of psychological flexibility (Blackledge, 2007; Hayes et al., 2012). Psychological flexibility has been defined as one's ability to remain in contact with the present moment, open to (i.e., accepting of) unpleasant private experiences, and engaged actions that reflect deeply held values (Harris, 2009; Hayes et al., 2012). Alternatively, psychological inflexibility is marked by avoidance of distressing emotions, preoccupation with judgments and/or memories, and disengagement with values-consistent action. ACT involves increasing flexible behavioral repertoires (discussed below) that may increase contact with personally reinforcing consequences while possibly reducing psychological suffering. Currently, most ACT protocols emphasize 6 core skills to increase functional (i.e., flexible) behavioral responses: present-moment awareness, cognitive defusion, self-as-context, values, acceptance, and committed action (Hayes et al., 2012).

Deficits in one of these ACT repertoires can be targeted individually or conjointly with other problematic repertoires in treatment by decreasing the response rate of relevant unhealthy behaviors while simultaneously increasing healthy and more functional behaviors. Present-

moment awareness is one's ability to attend to internal physical and psychological experiences, as well as external experiences within the environment at a given moment. Defusion is the ability to decrease the effect of one's cognitive experiences (i.e., thoughts) on emotional and behavioral responses by viewing cognition as an event rather than literally. Self-as-context is a perspective taking ability that allows an individual to recognize the constant stream of their internal and external experiences as being part of a larger context. Values represent the verbally mediated and intrinsically motivating ideals an individual holds. Acceptance is the willingness to openly experience unwanted thoughts, emotions, and sensations. Finally, committed action is the choice an individual makes to engage in any behavior(s) that is consistent with or allows them to engage with their values. While combinations of these processes are often targeted simultaneously in treatment, as stated above, the focus of the current paper is on defusion due to its direct relevance to problematic cognitions.

ACT is often referred to as a third-wave or new generation of CBT (Bluett, Homan, Morrison, Levin, & Twohig, 2014), a contention that is supported by empirical evidence indicating that ACT works through different mechanisms or processes than CBT (Hayes, Luoma, Bond, Masuda, & Lillis, 2006). Almost 200 randomized controlled trials have been published comparing ACT to other treatments for a host of physical and psychological concerns, and several meta-analyses have been published assessing its efficacy. For example, Hayes et al. (2006) found that ACT, and the processes that comprise ACT, have been shown to reduce fear, unwillingness to experience distressing emotions, and cognitive symptoms in female undergraduates with anxiety as assessed by the anxiety sensitivity index. A meta-analysis conducted by Bluett et al. (2014) found that ACT was statistically equivalent to CBT in the treatment of mixed anxiety disorders, social anxiety, and obsessive-compulsive disorder, while

also resulting in significant declines for symptoms of generalized anxiety disorder and panic disorder. Avdagic, Morrissey, and Boschen (2014) conducted a randomized controlled trial comparing ACT and CBT for the treatment of generalized anxiety disorder and results indicated that both treatments led to significant but comparable improvements for both conditions, although ACT demonstrated more reliable change at post-treatment than CBT. Furthermore, Hayes et al. (2006) noted two studies in which ACT produced larger decreases in math anxiety compared to systematic desensitization for individuals with higher experiential avoidance, as well as lower behavioral avoidance for a public speaking task, increased willingness to experience anxiety, and more time spent engaging in an exposure task compared to a group CBT condition.

Numerous other studies also indicate the effectiveness of ACT in the treatment of anxiety and other psychological disorders. One such study, conducted by Twohig et al. (2010), showed that ACT was more successful in the treatment of obsessive compulsive disorder than progressive relaxation training. Twohig et al. (2010) also found that ACT produced a significant reduction in depression symptoms for those participants reporting at least mild pre-treatment levels of depression. In a meta-analysis on the effectiveness of ACT compared to other treatments (e.g., medication management, CBT, 12-step programs) for substance use disorders, Lee, An, Levin, and Twohig (2015) found that ACT maintained a significant small to medium effect size over other treatments for various substance use disorders. In addition, ACT was shown to produce better long-term outcomes and lower rumination and depression scores, while CBT only reduced depression scores, in a college sample with depression (Zhao, Zhou, Liu, & Ran, 2013). Two other studies used an ACT protocol for the treatment of depression in veterans and results indicated that depression scores were significantly lower at the end of treatment for

younger and older veterans (Karlin et al., 2013; Walser, Karlin, Trockel, Mazina, & Taylor, 2013). In their meta-analysis Hayes et al. (2006) also found that ACT led to significant changes in social impairments, distress stemming from hallucinations, affective symptoms, and general functioning for patients suffering from psychosis or a mood disorder with psychotic features. In another study conducted by Wolitzky-Taylor, Arch, Rosenfield, and Craske (2012), results showed that CBT was more effective than ACT in the treatment of moderate baseline levels of anxiety, while ACT demonstrated better treatment outcomes than CBT for individuals with comorbid mood disorders. Arch, Wolitzky-Taylor, Eifert, and Craske (2012) took the comparison of ACT and CBT a step further by comparing proposed mediators (cognitive defusion and anxiety sensitivity, respectively) for both treatments in a clinical sample with anxiety disorders. Arch et al. (2012) found while both mediators demonstrated significant effect size increases for both treatments, ACT generated stronger treatment improvements and larger effect sizes for both mediators than CBT. Taken as a collection, these studies highlight the increasing evidence that ACT is at least as effective as CBT in the treatment of various psychological disorders and associated symptomology, and may potentially be more effective than CBT.

Although the six Hexaflex components of ACT all involve cognitive repertoires, defusion in particular bears a relatively explicit focus on problematic cognition, and in some ways may be said to complement the cognitive restructuring component of CBT. Defusion has been defined as the act of distinguishing between the semantic features of a thought and the sensory experiences that occur while having a thought (Blackledge & Drake, 2013). In other words, defusion is the ability to separate the meaning of one's thoughts from the emotional and physical responses that occur in response to a thought (Hayes et al., 2012), in order to reduce the influence one's

thoughts have on behavior (Arch et al., 2012). To highlight this process, imagine a scenario in which an individual is dealing with symptoms of depression following the loss of an intimate relationship. Such an individual may frequently experience problematic thoughts, such as "No one will ever love me again" or "I'm worthless." While experiencing these thoughts an individual will likely also experience intense emotional responses (e.g., sadness, depression, anxiety) and physiological responses associated with their emotions. In response to these thoughts, emotions, and physiological sensations an individual may begin engaging in dysfunctional behaviors such as withdrawing from all social contact and relationships. This is an illustration of fusion, or when problematic thoughts (and by extension emotions and sensations) rigidly dictate behavioral responses that contribute to the continuation and elaboration of the dysfunctional behaviors. If defusion is applied to this situation, the individual would be able to experience their problematic thoughts as just verbal stimuli without feeling compelled to control or change the thought or relevant emotional or physiological experiences associated with the thought. For example, through the use of various defusion exercises (some of which are described below) one will be able to experience the thought in several new ways including: (1) notice that s/he is experiencing a problematic thought, (2) notice the physical and sensory properties of that thought (e.g., how it sounds to say it, the letters and shapes made to write it, etc.), (3) become aware of the relationship the thought has with emotions and physiological responses in one's body, (4) notice how that thought interferes with valued actions, and (5) through these processes create "distance" between the thought and one's self thus allowing them to choose to engage in a more functional behavioral response.

Empirical support for defusion. Although the concept of defusion has been in the literature for only a couple decades, at least one technique has been around for a century that

captures the function of defusion, and is known as the word repetition technique (WRT; Titchener, 1916). When engaging in the WRT an individual is asked to quickly repeat a single word or short phrase aloud for 20 seconds or more, and while completing the exercise they are asked to notice how the literal meaning of the word moves outside of immediate awareness while the physical and sensory properties associated with the word become increasingly apparent (Blackledge, 2015; Blackledge & Drake, 2013; Hayes et al., 2012). Although Titchener (1916) first described this technique, in recent years a number of lab-based studies have utilized the WRT, and the most common WRT exercise in ACT is known as the "milk exercise." To elucidate the milk exercise, imagine a glass of milk sitting in on a table. This simple act of imagining a glass of milk likely brought to mind numerous qualities of milk, such as its color, taste, temperature, smell, consistency, the feel of a glass of milk in your hand, and possibly memories or emotional reactions one has had with milk previously (e.g., a memory of seeing a cow which produces milk). Interestingly, reading the sentence "imagine a glass of milk sitting on a table" is not the same thing as an actual glass of milk on an actual table, yet one's mind is able to draw forth this mental image when precipitated by a particular string of text. This is a frequently used strategy for teaching ACT clients about defusion: the inability to distinguish between one's thoughts (i.e., "think of a glass of milk") and the sensory or physical experiences that occur in addition to a thought. Next, the milk exercise requires an individual to quickly repeat the word "milk" for around 30 seconds and notice how the physical properties of saying the word become more salient while the literal meaning of the word dissipates over time. Through using the WRT or milk exercise clinicians are able to show clients that by focusing on the sensory details of a phrase or word that holds intense emotional weight they may be able to

weaken the emotional impact of those words, thus enabling clients to engage in values consistent behaviors that were previously adversely controlled by one's thoughts.

Numerous studies have examined the utility of defusion exercises in both lab- and clinically-based studies. For example, Masuda, Hayes, Sackett, and Twohig (2004) conducted a study comparing the WRT to thought control and distraction tasks to assess participants' believability and distress associated with negative self-referential thoughts. For the defusion condition, Masuda and colleagues provided a rationale and training in the WRT in accordance with a script for the technique in Hayes, Strosahl, and Wilson (1999), and then participants practiced the task using the milk exercise before employing the WRT for their own negative self-referential thoughts. In comparison, distraction condition participants were given an article on Japan to read during the intervention phase of the study, while the thought control participants were given a rationale focusing on the deleterious effects thoughts can have on one's life, and were then taught to use positive self-talk, positive imagery, and breathing exercises for controlling and preventing negative thoughts. To assess for the believability of and discomfort stemming from negative thoughts participants completed a 100-mm Likert-style visual analog scale for which participants rated the believability and discomfort for each thought ranging from 0 (not at all comfortable/believable) to 100 (very comfortable/believable) pre- and post-intervention. Findings indicted that the WRT successfully reduced distress and believability for negative self-referential thoughts in contrast to the other conditions, supporting the use of the WRT as a defusion technique.

As a follow-up to the above study, Masuda et al. (2009) conducted another analysis of the WRT by assessing whether altering the duration of the WRT affects the effectiveness of the task on decreasing believability and distress for self-referential thoughts. The first part of the study

had participants identify a single negative self-referential thought before rating the believability and discomfort for the thought using the 100-mnm Likert-style analog scale pre- and post-intervention. Participants were assigned to three distinct defusion conditions, each of which included a rationale and training in the WRT consistent with Masuda et al. (2004). Furthermore, the second and third conditions were asked to practice the WRT for a one-word negative self-referential thought for either 3 or 20 seconds. The results of this part of the study indicated that engaging in the WRT for a negative self-referential thought for either 3 or 20 seconds produced greater reductions in ratings of emotional distress than the condition that received only the rationale and training for the WRT, although there were no statistical differences in the use of the WRT for 3 or 20 seconds. Masuda et al. (2009) elaborated on the first part of the study in a second experiment in which participants were assigned to conditions that engaged in the WRT for either 1, 10, or 30 seconds after receiving a defusion rationale and training. Results of the second experiment showed that Likert-scale ratings for emotional distress were significantly lower for the 10 and 30 second conditions than the 1 second condition, with no statistical differences between 10 and 30 seconds. Additionally, Masuda and colleagues noted that ratings on the believability of negative self-referential thoughts continued to decline for up to 20 to 30 seconds of engaging in the WRT, showing that reductions in believability ratings persisted after emotional distress ratings had plateaued. Thus, consistent with Masuda et al. (2004), this study supports the use of the WRT for a minimum of 20 seconds in reducing believability and distress associated with negative self-referential thoughts.

Another study on the WRT has provided support for the use of defusion techniques in individuals experiencing increased depressive symptoms. Masuda Towhig, et al. (2010) assigned undergraduate students to a defusion intervention, an intervention based thought distraction

(training and practice in deliberately distracting one's self from a negative thought), or thought distraction control condition (reading an article on Japan for five minutes). Both intervention conditions received a clinical rationale and training in the assigned technique (WRT or thought distraction) before applying the techniques to the management of negative self-referential thoughts. Findings indicate that not only was the WRT superior to distraction conditions in reducing believability of and distress from one's negative thoughts, but the WRT successfully decreased distress and believability in participants with increased depressive symptomology. Masuda, Feinsten, Wendell, and Sheehan (2010) expanded on the findings of Masuda, Twohig, et al. (2010) by analyzing the effects of full or partial conditions of the WRT and thought distraction techniques. Partial conditions in this study included only a rationale and training in either defusion (WRT) or thought distraction (focusing on geometric shapes to distract one's self from negative thoughts). In contrast, full conditions also included an experiential exercise focusing on an identified one-word negative self-referential thought for each participant that lasted for 30 seconds (consistent with the findings of Masuda et al., 2009). Thus, four total conditions were included in the study. Findings showed that all conditions experienced significant declines in emotional distress and believability for the one-word thoughts following interventions. However, the full defusion condition achieved significantly greater reductions in comparison to the other three conditions for both believability and emotional distress. Therefore, Masuda, Feinstein, et al. (2010) showed that the addition of an experiential exercise in the training of defusion produces enhanced effects exceeding those produced by a rationale and training alone.

In addition to the many studies on the effectiveness of the WRT in lab-based settings, others studies have looked at different defusion techniques and/or applying defusion strategies to

more clinically relevant content. For example, Hinton and Gaynor (2010) assessed the efficacy of three therapy sessions of training defusion to undergraduates experiencing increased distress, dysphoria, and low self-esteem compared to a waitlist control condition. The defusion sessions focused on teaching participants how to notice their problematic thoughts as cognitive experiences that do not control their behavior. To achieve this, various defusion exercises and techniques were used in these sessions including the milk exercise, the contents on cards exercise (described in Hayes et al., 1999), and incorporating vocalizations strategies beyond word repetition. Results of the study indicated numerous beneficial effects for participants completing defusion training, such as decreased distress and dysphoria, increases in self-esteem, increased psychological flexibility, and better defused awareness of one's thoughts on a moment-to-moment basis. Moreover, these effects were replicated in the waitlist control patients who were given three sessions of defusion training after the waitlist period had completed, providing further evidence that defusion strategies, particularly verbalization techniques, are beneficial in the management of problematic cognitions.

Another examination of defusion conducted by Healy, Barnes-Holmes, Barnes-Holmes, and Keogh (2008) assessed the effects of having undergraduate students include the phrase "I am having the thought that" (a common defusion exercise; see Blackledge, 2015) before self-referential thoughts. The researchers posited that this exercise could reduce participants' believability and distress associated with negative self-referential thoughts (e.g., "I am stupid") by assisting participants in recognizing that thoughts lack innate truth and are simply creations of one's mind. Participants in the study were assigned to three conditions: (1) a pro-defusion condition emphasizing the helpfulness of defusion through a rationale and instructions of how to engage in defusion, (2) an anti-defusion condition in which participants received instructions

challenging the usefulness of defusion, and (3) a control condition. Participants in all conditions were shown 10 self-statements on a computer, each presented three different ways: (a) as a normally worded thought (e.g., "I am a bad person"), (b) as a defused thought (e.g., "I am having the thought that I am a bad person"), and (c) as an abnormally worded though (e.g., "I have a wooden chair and I am a bad person"). After being presented with each thought participants were asked to rate the believability of the thought, distress associated with the thought, and willingness to view and think about the thought further. Results showed that defused phrasing of thoughts led to increased willingness to view and think about the negative self-referential thoughts, and decreases in distress associated with such thoughts. Thus, this study indicates that simply rephrasing one's self-referential thoughts by including the phrasing "I am having the thought that" at the beginning can lead to a defusion effect of decreasing the believability and distress associated with problematic thoughts, while also increasing willingness to experience such thoughts.

Comparing defusion and cognitive restructuring. Some studies have demonstrated the efficacy of defusion in comparison to cognitive restructuring. Moffitt, Brinkworth, Noakes, and Mohr (2012) compared these strategies among a sample of self-identified chocolate cravers. Participants were randomly assigned to a waitlist control condition, a defusion condition, or cognitive restructuring condition, and all participants were given a bag of chocolates to carry with them for one week and asked to refrain from consuming any chocolate. Participants in the cognitive restructuring and defusion conditions completed standardized hour-long trainings for the respective techniques. Cognitive restructuring focused on the ability to challenge and alter one's thoughts by replacing them with more helpful thoughts, while defusion focused on creating distance between one's self from thoughts without challenging such thoughts. Findings show that

participants who engaged in defusion were over 3 times as likely to resist consuming chocolate than cognitive restructuring participants. In addition, individuals who had higher pre-treatment levels of distress performed significantly better in the defusion condition than the other two conditions, and participants self-reported that defusion was simpler to apply than cognitive restructuring and produced greater improvements in eating behaviors.

Deacon, Fawzy, Lickel, and Wolitzky-Taylor (2011) assessed the effects of defusion and cognitive restructuring on negative self-referential thoughts related to one's body shape in a sample of undergraduates. Participants initially completed the Body Shape Questionnaire to measure pre-treatment level of distress and concern with their body weight and shape, and then were asked to identify self-referential thoughts about and associated with being fat. Defusion condition participants followed the WRT protocol designed by Masuda et al. (2004), and after being taught the milk exercise were asked to employ the technique with each of their negative self-referential thoughts about being fat. In contrast, participants in the cognitive restructuring condition learned how to identify and challenge unrealistic thoughts regarding body image using a thought record to identify negative automatic thoughts and the situations in which they occur, evidence supporting and refuting a thought, and a more balanced thought based on the evidence. Results indicated that both conditions experienced comparable and significant decreases in distress and concern with body shape and weight, but the defusion condition participants responded more quickly to the intervention than cognitive restructuring participants. Moreover, defusion participants rated the thought of being fat as being less important post-treatment, whereas cognitive restructuring participants rated it as being more important post-treatment.

The collection of studies presented above support the use of defusion as an alternative treatment for problematic cognitions to cognitive restructuring. Multiple studies have shown that the WRT successfully reduces emotional distress and believability of negative thoughts (Masuda et al., 2004), including in people with increased levels of depression (Masuda, Twohig, et al., 2010) and individuals dealing with thoughts regarding body shape and weight (Deacon et al., 2010). In addition, the WRT has been found to produce the largest decreases in emotional distress and believability if the word repetition is engaged in for at least 10 and 20 seconds, respectively (Masuda et al., 2009). The inclusion of an experiential exercise during training for the WRT was also shown to lead to significantly stronger effects than simply providing a rationale (Masuda, Feinstein, et al., 2010). Beyond the WRT, defusion has also been shown to improve symptoms in people experiencing dysphoria, low self-esteem, and emotional distress (Hinton & Gaynor, 2010), and rephrasing one's thoughts by including the phrase "I'm having the thought that" effectively reduced distress and believability of negative self-referential thoughts while increasing willingness to experience such thoughts (Healy et al., 2008). When compared to cognitive restructuring for the management of food cravings, defusion exceeded the benefits of cognitive restructuring (Moffitt et al., 2012). Lastly, defusion produced treatment effects comparable to cognitive restructuring for problematic thoughts regarding one's body image, producing faster treatment effects and beneficial change in self-reported importance of such thoughts not generated by cognitive restructuring (Deacon et al., 2011).

As the above research shows, cognitive restructuring and defusion are two empirically supported approaches for the treatment of problematic cognitions. While the studies discussed above indicate that defusion may produce effects comparable to or exceeding cognitive restructuring, the research comparisons for cognitive restructuring and defusion are too few to

make a definitive conclusion. Moreover, there is ample evidence showing that cognitive restructuring is an efficacious strategy for treating problematic cognitions. Despite numerous studies supporting the effects of defusion on cognitions, the research base for defusion is still preliminary at this time due to the relative dearth of basic experimental studies assessing the core components and processes of defusion. For example, there has yet to be a detailed behavioral evaluation of defusion even though defusion effects are presumed to occur through the expansion of behavioral repertoires in response to one's problematic cognitions. To date, defusion studies have relied on self-report measures to assess for changes believed to be caused by defusion, showing that only cognitively-focused approaches have been used to measure the efficacy of defusion interventions. This approach to the assessing defusion is problematic because self-report measures have been shown to be vulnerable to personal biases and deliberate attempts to alter one's true responses (discussed below), further suggesting that the literature on defusion has been limited in the scope of its assessment of defusion treatment effects. Thus, future experimental studies on defusion should attempt to behaviorally assess the strength of learned verbal relations (i.e., fusion) pre- and post-administration of an intervention, thereby providing support or refutation of defusion techniques beyond those provided by self-report measures.

Measuring Cognition and Treatment Success with Self-Reports

Whether in randomized controlled trials assessing the effects of therapeutic techniques, or in lab-based studies conducting basic research on novel psychological constructs, researchers and clinicians alike extensively utilize self-report measures. The use of such measures is logical in that therapeutic success often requires participants to report, in a standardized way, whether or not they believe improvements have been made during the course of treatment. Thus, psychologists have devised countless measures of private experiences, including cognitions,

emotional distress, psychological symptoms, beliefs and attitudes, and personal assessments of one's previous behavioral experiences. Furthermore, self-report measures have an expansive history of rigorous methodological and psychometric evaluations, and provide a relatively quick and cost effective way of assessing treatment processes and outcomes. Nevertheless, there are multiple issues to consider when deciding if one should use self-report measures in a particular study, or if other forms of assessment may be more helpful in achieving one's research goals.

Despite their success and increasing use across the decades, self-report measures are susceptible to multiple issues that question the reliability and validity of such assessments. One such issue, referred to as the repeated measures effect (RME; Gilbert et al., 2002), occurs when scores on self-report measures of mood and affect decline due to repeated administrations and not a treatment effect. For example, Choquette and Hesselbrock (1987) noted that scores on the Beck Depression Inventory (BDI) and Zung Depression Scale significantly declined across three administration sessions in a sample of alcoholics. Without further analysis of the findings one would conclude that the decline in depression scores were due to a treatment effect, but Choquette and Hesselbrock (1987) reported that the deterioration of scores was not due to a treatment, but rather a seemingly natural decline in scores captured by the repeated administrations of the measures. This finding was replicated in Sharpe and Gilbert (1998) who administered several measures of mood (e.g., BDI, Profile of Mood States) either two or three times to participants who did not receive an intervention. Results of this study showed that mean scores for the BDI as well as the depression, tension, anger, fatigue, and confusion subscales of the Profile of Mood States (POMS) significantly declined across the multiple administrations. Subsequent to Sharpe and Gilbert (1998), Gilbert et al. (2002) noted a similar effect in a study assessing mood disturbance in a sample of individuals abstaining from smoking. Subjects were

administered multiple measures of mood states during a five-week baseline period. Results of this study indicated significant declines in scores on the BDI, Beck Anxiety Inventory, POMS, and Shiffman Withdrawal Questionnaire at the end of the baseline assessment period. Durham et al. (2002) provide a final example of the RME in a two-part study. In the first study researchers administered the Youth Outcome Questionnaire (Y-OC) to parents of elementary school students who were not receiving a treatment and noted that Y-OC scores significantly decreased across the duration of the study, with the largest decreases occurring for participants who were administered the Y-OC more often. During the second study undergraduates were administered the Outcome Questionnaire during a 9-week period during which participants did not receive a treatment, and results showed that scores on the measure significantly declined across the 9 weeks with the largest decrease occurring between weeks one and two.

Another common issue to consider with self-report measures is the possibility that socially desirable motivations are impacting responses. The concept of social desirability and response bias has been discussed within psychological literature for decades (e.g., Hanley, 1961), and refers to the ability for individuals completing self-reports to change their true responses to items, thus enabling them to engage in impression management when perceived by other people (Fleming, 2012). Not surprisingly, social desirability has been found to be present in studies looking at participants' physical appearance, psychological symptomology, attitudes and beliefs, behavior, and personality traits (Fleming, 2012; Vecchione, Dentale, Alessandri, & Barbaranelli, 2014). Additionally, research has found two separate forms of social desirability that can act jointly or independent of one another to alter item responses: self-deceptive enhancement (an almost narcissistic perspective of one's self), and impression management (altering responses to be perceived more favorably by others; Kuentzel, Henderson, & Melville, 2008). Interestingly,

individuals displaying higher trait levels of self-deception appear to make faster socially desirable responses, thus making it increasingly difficult to detect when such individuals are engaging in socially desirable behavior patterns (Holtgraves, 2004). Despite researchers' awareness that social desirability has affected self-reports and the construct validity underlying these measures, many psychologists continue to ignore or fail to address response bias concerns in psychological research (Ziegler, 2015).

Beyond the presence of the RME and socially desirable responding in self-report research, experimental error is an ever-present threat in the administration and scoring of psychological measures. For example, many measures of psychological symptomology and distress require the use of clinician judgment to determine if symptomology, or psychological disorders more broadly, are present in a given individual. Although psychologists (ostensibly) undergo rigorous training before being able to utilize such self-reports in clinical work or research, studies have shown that clinicians often fail to adhere to diagnostic criteria or guidelines for administering an assessment, leading to an over- or under-diagnosing of numerous individuals (Bruchmuller, Margraf, Suppiger, & Schneider, 2011). Therefore, the inclusion of self-report measures in psychology, especially measures of clinical symptomology, should be considered more cautiously due to potential concerns regarding their reliability and validity.

Considering the above evidence showing that self-report measures are vulnerable to various threats to their psychometric properties, such as the repeated measures effect, socially desirable responding, and experimental error, it is increasingly apparent that scores provided by self-report measures may be based on factors unrelated to therapeutic effects. For example, if the RME is active during a randomized controlled trial of a new therapy technique it will be nearly impossible to determine how much of a decline in scores is due to the RME or the therapy

technique independent of each other. Similarly, if researchers do not assess for social desirability, or guarantee the accuracy of experimenter administration and scoring, any conclusions drawn about a participant's self-report answers may be based on inaccurate scores that fail to capture one's "true" score. Consequently, future studies should expand their battery of assessments to include other methods of assessment beyond self-reports – even with the many successes such measures have provided to the field – that are not susceptible to the same issues as self-reports.

Assessing Cognition with Implicit Measures

Whereas explicit measures (i.e., self-reports) require introspection and overt responding to items presented to participants, implicit measures of cognition are based on the assumption that individuals are affected by previous experiences even when those experiences are not available for contemplation in conscious awareness (Greenwald & Banaji, 1995). In order to assess the influences of one's previous experiences on cognition and behavior implicit measures utilize reaction-time tasks to which participants respond as quickly as possible, thereby providing an estimate of the degree of strength for unconscious attitudes. Through the use of reaction-time tasks implicit measures are assumed to prevent participants from engaging in socially desirable responding to alter their responses. Additionally, participants completing implicit measures are frequently oblivious to the implicit attitude or cognition being assessed by the measure, further diminishing opportunities for participants to alter their true responses. Implicit measures also enable researchers to exert greater levels of control over the administration and scoring of the task compared to self-report measures, thus increasing the versatility of implicit measures beyond the capabilities of self-reports. For example, implicit measures have the capacity to be utilized as computerized analogue intervention delivery systems as De Young, Lavender,

Washington, Looby, and Anderson (2010) showed with the IAT. In short, implicit measures may be able to simultaneously act as a psychological intervention and measure of treatment effects. Below is a brief description of two implicit measures of cognition, and the research indicating their strengths and weakness in clinically relevant research.

Implicit association test. The implicit association test has been the most commonly used implicit measure over the last two decades, and requires participants to categorize two sets of concepts (e.g., white people and black people with good and bad) as quickly as possible. Nosek, Greenwald, and Banaji (2005) suggest that this methodology allows the IAT to measure the strength of associations between paired concepts by analyzing the response latency of items, and provide an estimate of the implicit cognitions that are influencing their behavior while completing the task. For example, if an individual can more quickly and accurately categorize "white people" and "good" then when "black people" and "good" are paired together this would indicate an implicit racial attitude that may not be captured on self-report measures due to socially desirable responding (i.e., most people resist admitting racial biases). The validity of the IAT has been demonstrated in multiple studies, such as in Lindgren et al. (2013) who used alcohol stimuli and approach or avoidance words and found that the IAT was able to predict past and future drinking behavior in college students. Similarly, Lindgren et al. (2013) found that a drinking identity IAT (i.e., relationship between self-identity words and alcohol stimuli) produced positive correlations with cravings, consumption, and alcohol related problems, as well as accounting for unique variance in drinking behavior after taking into account self-reported alcohol use. Furthermore, the IAT has been shown to discriminate between participants with a spider phobia, without a phobia, and those who enjoy spiders (Ellwart, Rinck, & Becker, 2006). In another IAT study, McConnell and Leibold (2001) noted that pro-black attitudes in the IAT

predicted more positive interactions between participants and black experimenters compared to white experimenters, but such a relationship was not found between behavior and self-report measures of racial attitudes. Taken together, these studies suggest that implicit measures of cognition may be better able to predict one's future behavior more successfully than self-report measures for socially sensitive attitudes and beliefs (Greenwald, Poehlman, Uhlmann, & Banaji, 2009).

Of interest to the current proposal are multiple studies showing that the IAT is able to detect changes in participants' cognition following a therapeutic intervention. For example, Teachman and Woody (2003) designed an IAT to assess negative attitudes toward spiders (e.g., "afraid" and "disgust") and results showed that after undergoing exposure therapy for spiders IAT scores were significantly decreased. Likewise, Blair, Ma, and Lenton (2001) found that attitudes toward gender stereotypes measured by the IAT decreased after participants were asked to imagine women with conflicting gender stereotypes. Ebert, Steffens, Stulpnagel, and Jelenec (2009) also found that one's implicit attitudes toward the self and others were altered after completing a single administration of IAT in which "self" and "positive" as well as "others" and "negative" were paired together.

Another important study for the current review was conducted by De Young et al. (2010) in which they compared the WRT to the Implicit Association Test (IAT) to assess the underlying mechanisms of defusion in word repetition strategies. The authors posited that because the IAT asks participants to sort categories/pairings of words as quickly as possible, participants are required to process the literal meaning of the words to perform the task. Thus, De Young et al. (2010) suggested that due to the repeated processing of the literal meaning of words in the IAT, the task may produce effects similar to the WRT while appearing more similar in function to

cognitive restructuring. The researchers had participants rate the emotional distress and believability associated with a list of negative self-referential words that were used in the IAT and WRT, and then assigned participants to one of five conditions: (1) control condition, (2) IAT without a rationale, (3) IAT with a rationale, (4) WRT without a rationale, and (5) WRT with a rationale. Participants received instructions for how to complete their respective tasks (IAT or WRT) and a rationale (if applicable) before completing either the IAT or WRT for their two highest rated words for distress and believability. Results showed that both tasks successfully decreased the believability and emotional distress in viewing or saying a negative self-referential word at post-treatment, but the WRT produced a greater decrease in emotional distress than the IAT suggesting that the act of repeating a word aloud (rather than repeatedly viewing the word on a computer screen) produces the greatest treatment effect.

While the IAT has demonstrated ample success in the measurement of supposed implicit cognition, there are concerns regarding the administration and scoring of the IAT that bring into question numerous conclusions that have been drawn from IAT studies. A significant issue of the IAT is that researchers report the strength of an association between different concepts, but are unable to provide an estimate of a single association independent of all other associations (De Houwer, 2002). This concern is exemplified in Greenwald, McGhee, and Schwartz (1998) who measured the association between the target concepts “flower” and “insect” and attribute concepts of “positive” and “negative.” Results of this study show that participants’ responses were faster when flowers and positive words and insects and negative words were paired together in comparison to opposite pairings (e.g., insect/positive and flower/negative). These findings indicate the strength of the relationship between flower and positive and negative words in relation to the relationship between insects and positive and negative words. Another way of

stating this is that the IAT is unable to explicate the strength of the relationship between flowers (a single target concept) and positive words (a single attribute concept), independent of its relationship with other attribute or target concepts (Greenwald et al., 1998). Another concern of the IAT is that it fails to directly assess a presumed implicit attitude because the task does not require participants to support or refute the relationship between target and attribute concepts (i.e., participants do not have to indicate that flowers *are* positive but simply sort these stimuli into the same category). Rather, researchers infer the existence of an implicit attitude based on how quickly participants sort paired concepts into a presumed higher-order implicit attitude category. Thus, the IAT has provided evidence supporting the use of implicit measures in clinically relevant research, but due to the subtlety inherent in the measurement of cognition – implicit or otherwise – an alternative measure of implicit cognition that addresses the concerns presented above for the IAT may be more pragmatic as an analogue therapeutic intervention and/or measure of treatment effects.

Implicit relational assessment procedure. The IRAP is a relatively new measure of implicit cognition that approaches cognition as a matter of stimulus relations rather than associations (Barnes-Holmes et al., 2006). The IRAP was developed based on behavior analytic principles espoused by Relational Frame Theory (RFT: Hayes, Barnes-Holmes, & Roche, 2001) in which human language and cognition is viewed as the primary stimulus control mechanism for human behavior due to the development of relational frames of reference (Hayes et al., 2012). Relational frames, or learned patterns of responding toward a stimulus based on its relation to another stimulus or group of stimuli, are based on three properties of increasing complexity: (1) mutual entailment, (2) combinatorial entailment, and (3) transformation of stimulus functions (Hayes et al., 2012). The first property, mutual entailment, refers to the bidirectional relationship

between two stimuli. For example, if an individual learns that the word "milk" refers to a glass with a specific white liquid in it, s/he will also be able to pick up the glass of liquid and call it milk without being taught to do this. Combinatorial entailment is the process of combining mutually entailed relational frames to form a relational network: if an individual also learns that the words "lemon" and "milk" are different, s/he will know without being taught that "lemon" and a glass of white liquid are different from each other. Transformation of stimulus functions occurs when psychologically meaningful applications of relational frames takes place, as evidenced by the knowledge that if an individual is vegan and is offered a cup of tea with the options of having milk or lemon in the tea, there will be a different transformation of function for milk and lemon due to the fact the milk comes from a cow (and is not consumed by vegans). Therefore, due to the IRAP's foundation in RFT principles it may be able to detect both directly learned relationships between stimuli (i.e., mutual entailment) and psychologically meaningful and novel applications of relational frames (i.e., transformation of stimulus functions).

Relational Frame Theory is also the underlying theory for Acceptance and Commitment Therapy and its associated processes (Hayes et al., 2012). Because of the relationship between RFT and ACT, one can explain the development of cognitive fusion through the framework of RFT to account for psychological distress stemming from problematic cognitions. First, the development of a growing relational network of words and objects begins immediately upon an individual acquiring language (e.g., "dog" is the same as the furry four-legged animal that barks). Next, over time the repeated exposure to specific relational frames (developed through mutual and combinatorial entailment) become associated with positive and negative emotional reactions. This process further elaborates the relational network through transformations of stimulus function that are either directly learned (e.g., happiness stemming from playing with a puppy, or

pain and anxiety after being bitten by a dog) or from arbitrarily derived relations (e.g., someone telling you "that dog is dangerous- stay away!"). This process occurs for countless objects and words in one's lifetime with little negative effect; however, psychological distress develops when the transformation of a stimulus function leads to the creation of rules that govern and restrict one's behavior, produce negative emotions and thoughts when not in the presence of the tangible stimulus, and prevents new learning opportunities from occurring. Using the example of the dog above, if bitten by a German shepherd a child may now generate the rule "that dog is dangerous" and then generalize this same rule to all other dogs, thus restricting the child's ability to perceive and respond flexibly to other dogs that are not dangerous. Moreover, this rule governed behavior could also contribute to negative emotions and problematic cognitions (e.g., rumination and worry) by simply hearing the word "dog" or seeing a picture of a dog, further restricting that child's behavior during subsequent interactions with dogs.

While RFT provided the foundation for the creation of ACT and the IRAP, the Relational Elaboration and Coherence (REC) model explains the process of completing the IRAP, and has been detailed by Barnes-Holmes, Barnes-Holmes, Stewart, and Boles (2010). During administration of the IRAP, participants view a sample stimulus (e.g., pictures of milk and lemons) at the top of the computer screen, a target stimulus (e.g., the words "milk," "lemon") below the sample stimulus, and two opposing response stimuli (e.g., "same" and "different") at either corner at the bottom of the computer screen. Over the duration of the procedure, each sample is paired with each target, thus generating four trial-types for an IRAP (e.g., picture of milk with "milk," picturing of milk with "lemon," picture of lemon with "lemon," and picture of lemon with "milk"). Participants must choose the correct response option based one of two rules regarding the relationship between the sample and target stimuli. One rule is hypothesized to be

consistent with participants' learning history, while the second rule is assumed to be inconsistent with their learning history. For example, using the stimuli above a consistent rule would be "Respond as if a picture of milk and the word 'milk' are the same" while the inconsistent rule would be "Respond as if a picture of milk and the word 'lemon' are the same." As participants complete the IRAP half of the trials require participants to respond in accordance with the consistent rule, with the remaining half of trials requiring correct responses based on the inconsistent rule. Thus, the REC model proposes that when participants respond to trials in a manner consistent with their learning history they will utilize relatively quick relational responses to choose the correct response option (i.e., respond quickly and accurately). In contrast, inconsistent trials should require a longer period of cognitive elaboration to make correct responses, leading to deviations in average response times between consistent and inconsistent trials that is referred to as the IRAP effect. The supposition that it is easier and faster for one to respond to stimulus pairings consistent with previous verbal learning history (and by extension one's implicit attitudes) is held by both IAT and IRAP research (Barnes-Holmes et al., 2006; Barnes-Holmes et al., 2010), and decreases the likelihood that an individual will attempt to engage in socially desirable responding due to the constraints of the IRAP.

Several studies highlight the advantages of the IRAP in comparison to the IAT. For example, Cullen, Barnes-Holmes, Barnes-Holmes, and Stewart (2009) conducted an ageism study in which participants completed an IRAP with "old people" and "young people" as sample stimuli, and various positive (e.g., happy, creative) and negative (e.g., sad, weary) words were used as target stimuli. Results of this study indicate that there was an initially strong pro-youth bias (i.e., young people are good), which was consistent with the findings of the similarly conducted IAT study by Dasgupta and Greenwald (2001) indicating participants were faster at

responding to young people/positive word pairings than for old people/positive word pairings. Each of these studies had a second part in which participants were shown exemplars of positively and negatively viewed old and young people, and results varied between the studies due to the inherent differences between the IAT and IRAP. Dasgupta and Greenwald (2001) found that after viewing pro-old exemplars participants responded faster to old people/positive word pairings and slower for young people/positive, indicating that implicit attitudes toward these groups of people had switched. In contrast, Cullen et al. (2009) found that after viewing the pro-old exemplars (e.g., Albert Einstein) the pro-young bias was reduced while the anti-old bias simultaneously moved to a more pro-old bias. These studies indicate the fundamental difference between IAT and IRAP research – the IRAP is capable of detecting changes in specific implicit attitudes for a single sample stimulus, whereas the IAT can only provide an estimate of one's bias for a sample stimulus in relation to another sample stimulus.

Roddy, Stewart, and Barnes-Holmes (2011) conducted another IAT/IRAP comparison study for weight-related attitudes in a sample of undergraduate students for the sample stimuli "thin" and "fat," and various positive and negative evaluations as target stimuli (e.g., "good" and "bad"). Results demonstrated that participants held stronger implicit pro-thin biases than indicated on self-report measures, with IAT and overall *D*-IRAP (an effect size estimate of implicit attitudes that takes into account response latencies across all trial-types) representing a significant pro-thin bias. However, the IRAP displayed differential findings on the trial-types in that anti-fat implicit attitudes were non-significant for both the Fat-Good and Fat-Bad trials, indicating that while IRAP participants exhibited pro-thin biases they did not display either a pro- or anti-fat bias (i.e., a neutral bias for fat). Much like Cullen et al. (2009), this study exemplifies the benefits of the IRAP over the IAT in that the IAT is unable to estimate the

strength of a single sample-target stimulus pairing, suggesting that the IRAP can more successfully measure the nuances of implicit cognition.

Several studies have also demonstrated the IRAP's ability to predict and measure clinically meaningful treatment outcomes. Carpenter, Martinez, Vadhan, Barnes-Holmes, and Nunes (2012) used the IRAP as a measure of treatment outcomes for 25 individuals with cocaine dependence who completed a 24-week outpatient treatment program. Participants were also administered a cocaine beliefs IRAP in which sample stimuli included "With Cocaine" and "No Cocaine," and positive and negative target stimuli included phrases such as "I am friendlier," "I am sexier," "I am paranoid," and "I am mean." Findings from the study indicate that participants displayed a significant negative *D*-IRAP score (i.e., an anti-cocaine bias), although there were variations in *D*-IRAP scores across participants. Furthermore, the IRAP demonstrated a moderate negative correlation with self-report treatment outcome measures that were completed at the beginning of the treatment program. These results suggest that individuals with a pro-cocaine bias, as measured by the IRAP, also had poorer treatment outcomes as evidenced by worse attendance for treatment and more frequent urine samples that tested positive for cocaine during the first half of the treatment program. In short, the IRAP was able to successfully predict behavioral treatment outcomes in a sample of cocaine-dependent individuals, and may be used in future research for other psychological disorders as a predictive measure of treatment outcomes.

The IRAP has also been shown to be successful in measuring fears related to specific stimuli, which is a feature of many psychological disorders (e.g., phobias). Nicholson and Barnes-Holmes (2012) used the IRAP to assess implicit biases towards spiders in individuals who were categorized with high or low spider fears. All participants completed a self-report measure of spider fear before completing an IRAP with sample stimuli reference approach or

avoidance of spiders ("Frightens me" or "I could approach"), with four pictures of spiders and four pictures of landscapes as target stimuli. After finishing the IRAP participants engaged in a behavioral approach task in which subjects approached a terrarium with a tarantula in it, and subjects were given scores based on how close they moved toward the terrarium. Findings of the study show that individuals with varying levels of spider fears (high or low) were successfully differentiated by the IRAP. Furthermore, the IRAP was able to predict subjects' performance on the behavioral approach task, further supporting the use of the IRAP as a potential predictor of treatment outcomes.

Another IRAP study examined emotional responses in a sample of normally and mild/moderately depressed participants as determined by the Depression, Anxiety and Stress Scale (DASS; Hussey & Barnes-Holmes, 2012). At the beginning of the study participants completed the 10-item Acceptance and Action Questionnaire-II (AAQ-II) as a measure of experiential avoidance and general psychological flexibility before being administered the IRAP, which was based on items from the DASS, to assess initial levels of emotional responses. The IRAP was comprised of sample stimuli with the phrasing "When things go well/badly" and positive and negatively worded target stimuli similar to "I feel happy/sad" so that participants responded to IRAP trials similar to "When things go badly…I feel sad." Next, participants completed a mood induction in which they were exposed to music to deliberately induce a sad mood, and this music was subsequently removed before participants were administered a second round of the AAQ-II and DASS-based IRAP. Hussey and Barnes-Holmes (2012) found that the AAQ-II and DASS were correlated with one another, showing that "normally" depressed subjects (i.e., within the normal range for depression on the DASS) exhibited higher psychological flexibility on the AAQ-II, while mild/moderately depressed subjects displayed

lower psychological flexibility. Interestingly, both groups of participants demonstrated a positive emotion response bias for the first IRAP administration, but after the mood induction IRAP scores remained stable for the normally depressed participants while mild/moderately depressed subjects evidenced significant declines in their emotional reaction. This is to say that the IRAP detected changes in emotional reactions following an intervention, indicating the IRAP can be used to measure changes in emotional and psychological symptomology.

In a study utilizing the IRAP to assess ACT processes, Hooper, Villatte, Neofotistou, and McHugh (2010) compared a thought suppression exercise to a mindfulness exercise in the treatment of experiential avoidance for problematic cognitions. Experiential avoidance (EA) is characterized by attempts to avoid or escape distressing internal experience (e.g., thoughts, emotions, physiological sensations), which is a lack of acceptance as described by Hayes et al. (2012). Baseline EA was assessed by administering the AAQ-II (presumed to be a self-report measure of EA) and an EA IRAP that included the sample stimuli "With negative emotions it is better that I" and "With negative emotions it is worse that I," and target stimuli such as "welcome them," "embrace them," "avoid them," and "reject them." Next, participants in the mindfulness condition completed a focused attention exercise for their distressing thoughts, whereas thought suppression participants were asked to suppress, ignore, or avoid thinking about distressing thoughts. Participants in both conditions were then exposed to an emotionally distressing image from the International Affective Picture Scale and asked to engage in the pertinent technique for their condition before completing the AAQ-II and EA IRAP again to assess for changes in EA. Hooper et al. (2010) found that participants in both conditions experienced mild decreases in EA on the AAQ-II, but mindfulness participants witnessed significant reductions in EA as measured by the IRAP whereas thought suppression participants

did not experience reductions in EA on the IRAP. Therefore, these results suggest that the IRAP is sensitive enough to detect changes in cognitions following interventions that self-report measures are unable to measure, supporting the use of the IRAP in the assessment of cognitive symptomology.

Kishita, Muto, Ohtsuki, and Barnes-Holmes (2014) assessed the effects of an ACT-based intervention using an IRAP. Participants with slightly elevated anxiety levels were randomized to either a defusion condition (engaging in the WRT), or a control condition (read an article about Japan). At the beginning of the study baseline levels of anxiety were assessed using an anxiety-relevant IRAP that included "Anxiety" and "Calmness" as sample stimuli, and aversive or non-aversive target stimuli such as "painful," "terrible," "pleasant," and "comfort." The premise behind using such an IRAP was that participants who respond in a fused manner to the IRAP will have rigid behavioral repertoires in response to both consistent and inconsistent block-types, whereas defused respondents will be able to respond more quickly and accurately to both block-types. Subjects then completed the relevant task(s) for their condition, which included learning how to engage in the WRT in the defusion condition, before completing the IRAP again as a post-intervention measure of anxiety. Results were generally consistent with the authors' supposition in that response latencies were significantly faster for consistent and inconsistent block-types in participants who completed a defusion intervention, but control condition participants demonstrated a small decrease in response times for inconsistent blocks only. Thus, consistent with Hooper et al. (2010), the IRAP successfully detected changes in implicit cognitions following an ACT-consistent intervention (i.e., defusion), which further supports the use of the IRAP as an assessment tool in clinically relevant research.

In a study similar to Hooper et al. (2010) and Kishita et al. (2014), Drake, Timko, and Luoma (2016) attempted to assess the implicit behavioral responses for experiential avoidance and acceptance in regard to positive and anxiety-relevant words using the IRAP. At the beginning of the study participants were asked to hold their breath for as long as they felt comfortable, and then completed an IRAP that included the sample stimuli "I am willing to have" and "I try to get rid of," and positive and anxiety-related target stimuli (e.g., happiness, relaxation, anxiety, fear). Following the IRAP several self-report measures of defusion, mindfulness, experiential avoidance, and distress were administered. The findings of this study indicated that while the breath holding task was not significantly correlated with the IRAP or any self-report measure, the IRAP and several self-reports demonstrated interesting correlations. For instance, AAQ-II scores were positively correlated with the "I try to get rid of anxiety" trial-type, indicating that more avoidance as measured by the IRAP is associated with lower AAQ-II scores. Similarly, lower scores on the State-Trait Anxiety Inventory (i.e., less distress) were more highly correlated with IRAP responses indicating high avoidance of anxiety, while higher anxiety scores were correlated in the expected direction with implicit biases for the trial-type "I try to get rid of happiness." In comparison, all five IRAP D-scores were correlated with a measure of defusion (Drexel Defusion Scale), with faster response latencies for each block-type of the IRAP being highly correlated with higher levels of defusion, but these correlations still indicated that higher defusion scores were associated with stronger implicit biases for avoiding anxiety. Thus, this study suggests the IRAP is more sensitive to fusion/defusion repertoires than avoidance or other behavioral repertoires due to the cognitive content represented in the IRAP, and the necessity of responding to the literal meaning of IRAP stimuli.

In another study of particular relevance to the current proposal, Ritzert, Forsyth, Berghoff, Barnes-Holmes, and Nicholson (2015) assessed the effects of a defusion intervention on behavioral repertoires as measured by the Implicit Relational Assessment Procedure (IRAP). Undergraduate participants with increased spider fears were randomized to a defusion condition (WRT training with an experiential exercise for the word "spider"), thought distraction condition (told to distract one's self from having a thought about spiders), or a control condition (reading an article). The IRAP, which showed pictures of spiders paired with phrases such as "disgusts me" or "I could approach," and a measure of thought believability were completed pre- and post-treatment. Results indicated that defusion, as expected, produced significant declines in thought believability compared to the control and distraction conditions. Additionally, defusion altered IRAP performance by producing IRAP effects close to zero at post-treatment (i.e., participants demonstrated more flexible responding when presented with pictures of spiders at the end of the study) while comparison conditions did not alter IRAP performance. Therefore, Ritzert et al. (2015) demonstrated that a defusion intervention can exhibit a direct influence on IRAP performance, suggesting defusion can impact implicit cognitions for clinically relevant content and that the IRAP is capable of measuring such changes.

Taken together, these IRAP studies provide consistent support for the use of the IRAP as (a) an alternative for assessing cognitive experiences beyond self-reports, (b) as a successful alternative to the IAT in the measurement of implicit cognitions, (c) as a tool to detect nuanced changes in implicit cognitions following interventions, and (d) as a predictor of behavioral treatment outcomes. As the above studies show, the IRAP is a behavioral measure that has been shown to effectively assess implicit cognitions regarding participants' body weight (Roddy et al., 2011), young and old people (Cullen et al., 2009), cocaine (Carpenter et al., 2012), spiders

(Nicholson & Barnes-Holmes, 2012; Ritzert et al., 2015), emotional reactivity (Hussey & Barnes-Holmes, 2012), and anxiety (Kishita et al., 2014). Furthermore, the IRAP has been empirically supported as a predictor of behavioral treatment outcomes for individuals with cocaine dependence undergoing treatment (Carpenter et al., 2012) and on a behavioral approach task for a tarantula in people with heightened fear of spiders (Nicholson & Barnes-Holmes, 2012). Most importantly for the current proposal, the IRAP has demonstrated efficacy in measuring changes in implicit cognition following a mindfulness intervention to reduce experiential avoidance (Hooper et al., 2010), a defusion intervention for individuals with higher anxiety levels (Kishita et al., 2014), and a defusion intervention for spiders in participants with increased spider fears (Ritzert et al., 2015). Despite the IRAP's success in measuring the effects of defusion across several studies, there has not been a study assessing the basic foundations and processes of defusion. By taking a bottom-up, basic experimental approach to analyzing the effects of defusion, researchers and clinicians may be able to show further empirical support for the treatment of problematic cognitions, as well as potentially enhancing the effects of such interventions in the future. Therefore, the IRAP appears to be a reasonable addition to research studies assessing the effects of ACT-based interventions (e.g., defusion) on one's cognition due to the shared underlying principles of Relational Frame Theory within ACT and the IRAP.

The Current Proposal

Of the seemingly countless psychological disorders discussed in the Diagnostic and Statistical Manual of Mental Disorders-Fifth Edition (American Psychiatric Association, 2013), problematic cognitions are a defining feature of a large portion of disorders. Not surprisingly, clinicians have developed many treatments targeting problematic cognitions through various means. Both CBT and ACT have been shown to be effective in the treatment of various

psychological disorders, and studies comparing the effectiveness of the treatments tend to find that CBT and ACT are comparable to one another (e.g., Arch et al., 2012; Wolitzky-Taylor et al., 2012; Zhao et al., 2013) despite disparate proposed mechanisms of change and differential treatment outcomes. For the treatment of problematic cognitions CBT utilizes cognitive restructuring to challenge the veracity of dysfunctional thoughts, while ACT incorporates defusion to reduce the impact of the literal meaning of thoughts on behavior. Cognitive restructuring has been empirically tested in a variety of ways over the last 50 years, and the findings as a whole have largely supported this strategy, but with inconsistencies at times. For example, several studies have found that the addition of cognitive restructuring to other treatments (e.g., exposure therapy) has produced better treatment effects than conditions not receiving cognitive restructuring (e.g, Bryant et al., 2003; 2008; Mattick & Peters, 1998), but others show cognitive restructuring provides no additive treatment effects (e.g., Salaberria & Echeburua, 1998; Scholing & Emmelkamp, 1996). In contrast, defusion has been demonstrated to produce beneficial treatment effects through the use of various defusion interventions, including the word repetition technique (e.g., Masuda et al., 2004; 2009; 2010) and a thought rephrasing strategy (e.g., Healy et al., 2008). However, the body of ACT research studies continues to grow, and dismantling studies will provide clarifications on the influence of defusion as a component of treatment for problematic cognitions. In addition, multiple direct comparisons of defusion and cognitive restructuring have been completed showing that defusion can produce faster treatment responses (Deacon et al., 2011), significantly increase the likelihood to resist eating a desired food compared to cognitive restructuring (Moffitt et al., 2012), and has been rated as simpler to use than cognitive restructuring (Moffitt et al., 2012). Thus, ACT and

defusion appear to be gaining empirical support as a viable alternative treatment to cognitive restructuring for problematic cognitions, and should be considered for further empirical testing.

Throughout the period of empirical investigation of psychological therapies, self-report measures have been the most widely used method of assessing treatment success; however, self-reports have been shown to be vulnerable to socially desirable responding (Fleming, 2012; Kuentzel et al., 2008) and the repeated measures effect (e.g., Gilbert et al., 2002; Sharpe & Gilbert, 1998). In contrast, implicit measures of cognition, such as the IRAP, are behavioral tasks that measure cognitive processes by having participants respond as quickly and accurately as possible to presented stimuli. Since its inception, the IRAP has been shown to measure implicit cognition for a host of psychological concerns, and has successfully detected changes following therapeutic interventions in experiential avoidance (Hooper et al., 2010) and emotional reactivity in depression (Hussey & Barnes-Holmes, 2012). Furthermore, multiple studies have demonstrated that implicit cognition can be influenced by training in defusion, as evidenced by significant changes in *D*-IRAP scores and response latencies following defusion interventions (Kishita et al., 2014; Ritzert et al., 2015). The IRAP has also established good convergent validity with well-validated self-report measures of cognition, and discriminant validity with other less reputable self-report measures that are more susceptible to social desirability and the RME (see Golijani-Moghaddam, Hart, & Dawson, 2013). Additionally, the IRAP has exhibited generally acceptable to good internal consistency and test-retest reliability for a behavioral measure (Golijani-Moghaddam et al., 2013), providing additional support for the use of the IRAP as an assessment tool for implicit cognition in clinically relevant psychological research.

The current proposal identifies the growing body of empirical support for the use of cognitive defusion in the treatment of problematic cognitions (e.g., Masuda et al., 2004; 2009;

Hinton & Gaynor, 2010), while also recognizing concerns regarding the use of self-reports to assess the effects of therapeutic interventions as discussed previously. In addition, there is ample evidence to support the use of implicit measures in psychological research broadly (Barnes-Holmes et al., 2006), as well as in clinical research (e.g., Carpenter et al., 2012; Nicholson & Barnes-Holmes, 2012). Existing IRAP studies on defusion have focused on complex stimulus relations, and as a result defusion effects in these studies have been isolated to specific aspects of IRAP data. The complexity of these paradigms has resulted in relatively unclear and unpredictable defusion effects within the IRAP across studies. In an attempt to conduct a more basic test of the effects of a defusion intervention on basic learned verbal relations, the current study utilized the empirically supported WRT on a basic stimulus network involving the words milk and lemon. The stimuli "milk" and "lemon" were chosen for the current study because they are the focus of common, well-established interventions in ACT protocols and because they offer simple stimulus products for inclusion in the IRAP. This simple paradigm may provide a foundation for further studies that could carefully and systematically dismantle the effects of the WRT on a network of stimulus relations using a behavioral measure. Beyond replicating findings that the IRAP is sensitive to changes in implicit cognition following a defusion intervention (e.g., Kishita et al., 2014; Ritzert et al., 2015), this study may provide a means of more thoroughly assessing the effects of defusion via principles of RFT. Specifically, the current study attempted to accomplish two goals:

- Replicate previous research showing that the IRAP is sensitive to changes in implicit cognition following a therapeutic intervention (e.g., Hooper et al., 2010), specifically after participants receive training and an experiential exercise in cognitive defusion (e.g.

Kishita et al., 2014; Ritzert et al., 2015). Based on this hypothesis, the following results are predicted in conditions receiving defusion:

    - Significant differences in *D*-IRAP scores for the two trial-types containing the word "milk" between the control condition and a defusion condition that completes the WRT for the word "milk"
    - Significant differences in *D*-IRAP scores for trials containing the word "milk" and the word "lemon" between the control condition and a defusion condition that completes the WRT for both "milk" and "lemon"

- Replicate previous research showing that a defusion intervention impacts IRAP performance by reducing response latencies for both consistent and inconsistent block-types (e.g., Drake et al., 2016; Kishita et al., 2014) for only IRAP stimuli that have undergone a defusion intervention. Specifically, this hypothesis predicts:
    - Significant differences in response latency for consistent and inconsistent trials containing the word "milk" between the control condition and a condition receiving defusion for the word "milk"
    - Significant differences in response latency for consistent and inconsistent trials containing either the word "milk" or "lemon" between the control condition and a condition receiving defusion for the word "milk" and "lemon"

CHAPTER 2

METHODS

Participants

One-hundred and twenty-two undergraduates (74 female, 48 male) at Southern Illinois University received course credit for participation in the study (Mean age = 19.76, SD = 1.71; one participant reported an inaccurate age). Of these 122 participants, 11 failed to complete the IRAP at a predetermined criterion level across all blocks, and were excluded from analyses. Thus, 111 participants (68 female, 43 male; Mean age = 19.79, SD = 1.77) were included in the analyses detailed below. Of the participants included in analyses, 52 (46.8%) were freshmen and 105 (94.6%) reported English as their first language. Fifty-three (47.7%) subjects identified as being White or Caucasian, 35 (31.5%) identified as Black or African-American, 7 (6.3%) reported being Hispanic/Latino, 5 (4.5%) identified as Asian, and 9 (8.1%) identified as multiracial. Seventy-six (68.5%) reported being Christian, while 19 (17.1%) reported being (Agnostic). Eighty-six (77.5%) denied have previous therapy experience, 42 (37.8%) reported an SES of $25,000 or less, 25 (22.5%) reported an SES of $25,001-$50,000, 22 (19.8%) reported an SES of $50,001-$75,000, and 22 (19.8%) reported an SES above $75,000. A total of 29 participants were included in the control math condition, while 26 were in the control defusion condition, 27 in the half defusion, and 29 were included in the full defusion condition.

The current study asked participants to complete a set of questionnaires and a computer task during which participants responded to words and images related to milk and lemons. During completion of certain questionnaires regarding participants' psychological functioning, it was possible that individuals could have experienced varying levels of discomfort or distress (e.g., cognitive or physiological experiences in response to a question on a questionnaire), but no

participant reported experiencing emotional or psychological distress to an experimenter. Moreover, previous research has not shown evidence of deleterious long-term effects from using the IRAP (e.g., Hayes et al., 2006). In addition, participants were asked to rate various words after completing the IRAP that could represent fused or defused responses to "milk" and "lemon."

Once participants arrived, they were guided through an informed consent process, during which all participants were informed that 1) participation in the study was voluntary, 2) participants could choose to not complete any part of the study without forfeiting their credit for participating in the study, and 3) informed of potential risks for participating (e.g., emotional or psychological distress during completion of questionnaires). Before completing questionnaires, participants were briefly informed on the content of the various questionnaires, and were requested to complete all questions honestly to ensure the accuracy and validity of their responses. Each participant completed the study individually as a means of limiting potential distractions from other participants, as well as a means of increasing the experimental control of the study. After participants completed all aspects of the study they were given a debriefing form that described the purpose of the study and the researcher was available at this time to answer any questions participants had regarding the study. If a researcher had observed a participant experiencing distress while completing the study the researcher would have reiterated that participation is voluntary and that the participant could refrain from completing the study without penalty, but as stated above no participant reported or was observed experiencing distress. These guidelines, including the informed consent, debriefing form, individual completion of the study, and voluntary participation, allowed for enhanced protection of each participant's psychological and emotional well-being throughout the course of the study.

Design

The current stud utilized an experimental design in which participants were randomly assigned to one of four conditions: 1) a control group, 2) control defusion group, 3) half defusion condition (targeting only the word milk), and 4) full defusion condition (targeting the words milk and lemon). Participants in all conditions received the same measures, and varied only on their level/kind of training with defusion. Additionally, participants' baseline levels of defusion and psychological functioning were assessed as a means of checking for pre-existing differences between conditions and as a potential covariate in analyses for implicit attitudes towards milk and lemons. Successful acquisition and application of defusion were determined by comparing IRAP scores in the defusion conditions in relation to the control math condition, with the expectation that participants trained to defuse from lemons and/or milk would show statistically significant differences in IRAP D-scores. Thus, D-scores were used as the primary outcome variable.

Measures

Self-Report Measures

Demographic Questionnaire. Participants completed a series of demographic questions regarding their age, country of origin, ethnicity, sexual identity, gender, education level, membership in a sorority or fraternity, declared major, English as a first language, socioeconomic status, religious affiliation, political affiliation, and if they experience any food allergies. These questions allowed the researcher to assess for balance across experimental conditions based on random assignment. The demographics questionnaire also included a single item regarding participants' previous therapy experience, if applicable, to allow the researcher to

control for possible treatment effects that could affect the current study. See Appendix A for the demographics questionnaire.

Cognitive Fusion Questionnaire. The Cognitive Fusion Questionnaire (CFQ; $\alpha = .91$; see Appendix B) is a 7-item measure of cognitive defusion based on ACT and RFT that uses a 7-point Likert scale ranging from 1 ("*never true*") to 7 ("*always true*"). Scores on the CFQ range from 7 to 49, with higher scores indicating higher levels of fusion (lack of defusion). Gillanders et al. (2014) describe a multi-part study on the creation and psychometric validation of the CFQ in which experts in ACT and RFT developed a list of questions to assess defusion and fusion, and then rated these questions based on how well they represent aspects of defusion to begin with a list of 42 questions. In the first study, comprised of various samples of participants, the researchers identified a grouping of 7 items representative of fusion, and a confirmatory factor analysis was used in study two to confirm that this factor of 7 items best represents cognitive fusion. In study three the validity of the CFQ was assessed in comparison to various other measures of psychological functioning, and demonstrated strong convergent validity (rs $\geq .69$) with other measures of psychological distress such as the AAQ-II, BDI-II, Center for Epidemiological Studies Depression Scale, and Ruminative Response Style Questionnaire. The CFQ also demonstrated expected discriminant validity (rs $\leq -.50$) with the Southampton Mindfulness Scale, Five Facets Mindfulness Questionnaire, and Kentucky Inventory of Mindfulness Skills. Furthermore, the CFQ demonstrated a test-retest reliability of .81, and an internal consistency across the various samples of no less than .88. Thus, the CFQ appears to be a valid and reliable measure of psychological distress, consistent with ACT and RFT, and was utilized in this study to assess for comparability of scores between conditions.

Depression Anxiety and Stress Scale 21. The Depression Anxiety and Stress Scale 21 (DASS 21; $\alpha = .95$; see Appendix C) is a 21-item measure of the multiple dimensions of anxiety, depression, and stress that uses a 4-point Likert scale ranging from 0 (*never*) to 3 (*almost always*). The 21 items are split into three separate scales (depression, anxiety, and stress), and three scales scores are calculated that range from 0 to 21 with higher scores indicating more severe levels of symptomology. Lovibond and Lovibond (1995) detail the initial validation of the DASS in which 42 questions for negative emotional symptoms were administered to 717 undergraduates along with the Beck Anxiety Inventory (BAI) and BDI. Findings of the study showed that the DASS can successfully differentiate between the three forms of distress it assesses and is highly correlated with the BAI and BDI (rs $\geq .74$) as expected, demonstrating good convergent and divergent validity. Osman et al. (2012) conducted two studies assessing the reliability and validity of the DASS in which 887 and 410 undergraduate students, respectively, were administered the DASS with other symptom measures. Results of this study show that DASS items load onto three factors as expected, internal consistency estimates were good for each of the scales (all $\alpha \geq .81$), and the DASS was highly correlated with the BDI-II, a measured of mixed anxiety and depression, and a measure of perceived stress. However, Osman et al. (2012) noted that a DASS total scale score (incorporating all subscales) may be beneficial because many of the items strongly loaded onto a factor of general distress. Nonetheless, the DASS appears to be a reliable and valid measure of symptoms of depression, anxiety, stress, and general distress. The DASS was used in the current study to assess for comparability of scores between conditions.

Food Attitudes Questionnaire. This questionnaire was created to assess participant attitudes toward various food-related items. It consists of four items that ask participants to

indicate their frequency and level of enjoyment in consuming four drinks (milk, water, coffee, and pop/soda) and four foods (apples, lemons/lemon flavored foods, oranges, and bananas). Items assessing frequency of ingesting each food ask participants to indicate how many days of the week each food or drink is consumed, while items measuring one's enjoyment in eating or drinking a food uses a 7-point Likert scale ranging from -3 (*not at all*) to 3 (*very much*). After participants completed this measure, the experimenter assessed participants' responses to determine if this measure could be included as a covariate. See Appendix D for the full measure.

Milk/Lemon Defusion Measure. This measure was created for the purposes of the current proposal to assess if the defusion intervention(s) completed by the half and full defusion conditions had an effect on participants' self-reported reactions to viewing the words milk and lemon. The Milk/Lemon Defusion Measure consists of two items that ask participants to rate the saliency of various reactions to the words "milk" and "lemon." Participants rated four qualities relevant to milk (white, creamy, cow, and dairy) and lemon (yellow, fruit, citrus, and juicy), as well as four defused reactions participants may have upon reading "milk" and "lemon" (word, sound, text, and symbol) on an 8-point Likert scale ranging from 0 (*not at all*) to 7 (*extremely*). Participants completed this measure to assess for effects of defusion following an intervention and completion of an implicit measure. See Appendix E for the full measure.

Multidimensional Psychological Flexibility Inventory. The Multidimensional Psychological Flexibility Inventory (MPFI; $\alpha = .90$; see Appendix F) is a 60-item measure of psychological flexibility consistent with Acceptance and Commitment Therapy. The MPFI was created by Rolffs, Rogge, and Wilson (2016) to be a more comprehensive measure of the flexible and inflexible behavioral repertoires within ACT, whereas other ACT consistent measures tend to focus on a minimal range of flexible/inflexible behaviors. To achieve this goal, Rolffs et al.

(2016) conducted an initial study in which 372 participants were administered 494 items, which included items from 22 existing ACT-relevant measures and additional 84 items created by the researchers. The 494 items were grouped into 12 dimensions of flexibility/inflexibility, and each grouping of items was subjected to an exploratory factor analysis to identify the items that most strongly fit each dimension, leaving 214 items within the measure. The authors conducted a second study in which these 214 items, as well as 74 additional items created by the authors, were administered to 2,150 participants, thus allowing for the use of item response theory (IRT) to further refine the MPFI items. Following IRT analyses, 60 items (five for each dimension) were identified for the final MPFI with 8 of the dimensions capturing equivalent or higher levels of variance than comparison measures relevant to each dimension, and all of the dimensions demonstrated good internal consistency (all $\alpha s \geq .85$). Rolffs et al. (2016) also reported good convergent validity between the MPFI and other well-known ACT-relevant measures (e.g., Acceptance and Action Questionnaire-II and Multidimensional Experiential Avoidance Questionnaire) Five Facet Mindfulness Questionnaire, and Self-Compassion Scale), as well as divergent validity with conceptually distinct scales. Thus, the MPFI is a reliable and valid measure of psychological flexibility. The MPFI was utilized in the current study to assess for comparability between conditions.

Implicit Relational Assessment Procedure

The IRAP is a reaction-time task based on RFT that measures implicit cognition by having participants respond to visual stimuli (words or images) as quickly as possible to determine if the relationship between presented stimuli is consistent or inconsistent with one's learning history. Barnes-Holmes et al. (2006) detailed the IRAP and its functions, which asks participants to respond to a series of blocks containing randomly administered trials that present

a participant with a sample stimulus at the top of the computer screen, a target stimulus below the sample stimulus, and response stimuli at the bottom of the screen. For the current study, subjects saw either one of six pictures of milk or one of six pictures of a lemon as the sample stimulus, the word "milk" or "lemon" as the target stimulus, and the words "same" or "different" as response options. Upon being presented with these stimuli participants were asked to choose one of the response options by pressing the "d" or "k" key on the computer keyboard, where "d" was associated with the response in the bottom left of the screen and "k" was associated with the bottom right response option on the screen, although the actual response options (same and different) switched sides randomly. When participants chose the correct response option the computer screen was cleared (i.e., a blank white screen is displayed) for 400 ms before another trial was presented. In contrast, an incorrect response led to a red "X" appearing in the middle of the screen indicating that the participant had chosen the wrong response, and they had to choose the correct response before the next trial was administered.

A correct answer on the IRAP was determined by the block-type a participant completed, with even-numbered blocks requiring one set of responses on accompanying trials, and odd-numbered blocks requiring a different set of responses on trials. Thus, a block-pair consisted of a sequential pair of an odd- and even-numbered block (e.g., blocks one and two, blocks three and four, etc.). Each block-pair consisted of two block-types, consistent and inconsistent, where the consistent blocks required participants to respond to stimuli in a manner that is believed to be consistent with their previous learning history (e.g., a picture of milk and the word "milk" require the selection of "same"), and inconsistent blocks require participants to respond in a way that is inconsistent with previous learning experiences (e.g., a picture of a lemon and the word "milk" require the selection of "same"). Each participant was able to complete up to three

practice block-pairs to learn how to respond to the trials, and then subjects were required to complete three additional test block-pairs for the IRAP. This IRAP was designed to assess the strength of participants' implicit attitudes between each possible combination of images and words, resulting in four trial-types (a pairing of a sample stimulus with a target stimulus). The four trial-types include image of milk + "milk", image of milk + "lemon", image of lemon + "lemon", and image of lemon + "milk". These trial-types were randomly ordered during a given block with no more than two of the same trial-type presented in a row (see Appendix G for examples of trial-types as they appear in the IRAP).

Before beginning the IRAP researchers presented participants with instructions similar to those of Drake, Seymour, and Habib (2016) in which the procedures of the IRAP are clarified in a step-by-step manner to the participants. While administering the instructions, participants were presented with figures showing the four trial-types and were guided through a description of how to correctly respond to each trial-type in accordance with the block-type being presented. All participants received the consistent block-type first, and before beginning a block-type participants were shown a computer screen indicating the rule to follow (described above) for the subsequent block-type. As participants completed the IRAP they were required to respond with at least 78% accuracy across the 24 trials of any given block-type, and respond with a median latency of less than 2000 ms for the block. If a participant took longer than 2000 ms to respond to an individual trial the symbol "!" appeared in red text in the middle of the computer screen to prompt the participant to respond more quickly on remaining trials of the block. As stated above, participants completed up to three practice block-pairs before three test block-pairs were presented, and if a participant's accuracy or latency fell below acceptable standards the researcher immediately offered feedback and guidance to the participant in an attempt to increase

adherence to IRAP criteria. To calculate each participant's implicit biases for milk and lemons the response latencies in the IRAP were transformed into *D*-IRAP scores following the procedure described by Vahey, Barnes-Holmes, Barnes-Holmes, and Stewart (2009), which were used in various analyses (see Appendix H for procedures to calculate *D*-IRAP scores).

Over the last decade of the IRAPs use in implicit research, numerous studies have assessed the reliability and validity of the IRAP. Golijani-Moghaddam et al. (2013) conducted a meta-analysis of available IRAP research and found the internal consistency of the IRAP to be .81 when latency criteria was 2000 ms. Several studies have also assessed various forms of validity with the IRAP. For example, Golijani-Moghaddam et al. (2013) reviewed available research on the convergent validity of the IRAP with the IAT and found that the two implicit measures often show weak correlations with each other across studies, but as the content (stimuli) of the IRAP and IAT becomes more similar their convergent validity increases. Additionally, the IRAP has demonstrated good convergent validity with several well-validated self-report measures, including clinically-relevant content such as spider fears (Nicholson & Barnes-Holmes, 2012; Ritzert et al., 2015) and obsessive-compulsive tendencies (Nicholson & Barnes-Holmes, 2012). Furthermore, the IRAP has demonstrated predictive validity in multiple studies, such as Carpenter et al. (2012) that showed the IRAP was able to predict treatment success in cocaine dependent participants. Moreover, the IRAP has shown the ability to accurately discriminate between different groups of people based on their IRAP performance, such as meat-eaters and vegetarians, individuals of various social groups/classes, convicted child sex-offenders from non-offenders, "normal" and "mild/moderate" depressed groups, and level of spider fears (see Golijani-Moghaddam et al., 2013). Thus, based on the results presented by

Golijani-Moghaddam et al. (2013) and other studies described above, the IRAP demonstrates acceptable reliability and validity as a behavioral measure of implicit cognition.

Procedure

Group Assignment

Participants voluntarily decided to participate in the current study through an internet-based research participation program associated with their *Introduction to Psychology* course (Sona-Systems), and each participant received partial credit for their class after completing the study. Upon signing up for research participation in the study, participants were provided with the time and place of the study, while all other information (purpose of study, principal researcher, other participants partaking in study, etc.) was withheld to improve the randomization of participants to study conditions. After arriving to the study participants were directed to sit at a specific computer and turn off their electronic devices to further decrease potential distractions. Next, participants were guided through an informed consent document (see Appendix I), and the researcher was available to answer any participant questions at this time. Before administering any questionnaires or the IRAP, researchers ensured that participants were aware of all possible risks and benefits for their participation in the study, and were informed that they could withdraw from the study without penalty at any time. Participants were randomly assigned to one of the four conditions. Randomization was conducted upon the arrival of participants for the study. Each room in which the study was conducted had a separate form to track participant completion of each component of the study, and on this form participant conditions were randomized through a numbering system so that all four conditions received one out of every four participants that arrived to the study, relative to each individual room. The experimenter tracked the number of participants in each condition across all rooms utilized for the study to ensure that

participants were being distributed evenly across the four conditions. Moreover, the experimenter checked each condition for demographic breakdowns after every 50 participants to determine if participants were being randomized across conditions based on demographic information as well (e.g., race, gender, age). After beginning the study and being assigned to a condition participants completed a series of self-reports. Following completion of the self-reports, participants completed an exercise based on their condition (see below), and then completed a single milk/lemon IRAP. Following administration of the IRAP, all participants were given the Milk-Lemon Defusion Measure, and then given a debriefing form at the conclusion of the study. A flowchart of experimental procedures is displayed in Appendix J.

Conditions

Participants were randomly assigned to one of four conditions, including a control math (CM) condition, a control defusion (CD) condition, a half defusion (HD) condition, and a full defusion (FD) condition. All participants first completed a series of questionnaires assessing psychological functioning and attitudes toward various foods and drinks (including milk and lemons). Next, CM participants completed a series of simple math problems for a length of time equivalent to training and administering defusion in the FD condition (four minutes). Following this, subjects were instructed on how to complete the IRAP, and then administered the milk/lemon IRAP to assess relational repertoires for milk and lemons. After completing the IRAP, CM participants completed the Milk/Lemon Defusion measure to assess whether participants displayed more defused attitudes towards “milk” and “lemon” after completing the IRAP. Lastly, participants were given a debriefing form (see Appendix K) detailing the purpose of the study, and the researcher was available to answer any questions. The CD, HD, and FD conditions followed the same procedures as the CM condition, except that all defusion-based

conditions heard a rationale for defusion and then engaged in a defusion exercise (described below) instead of completing simple math problems. Subjects in the CD underwent defusion training for stimuli not presented in the IRAP ("car" and "rabbit"), and completed the word repetition training for both of these words. Subjects in the HD condition completed training for defusion as it relates only to "milk," and completed the word repetition exercise accordingly. In comparison, participants in the FD condition complete the word repetition for the words "milk" and "lemon." Following the word repetition exercise, participants in all defusion conditions were guided through the remaining procedures of the study consistent with the CM condition described above.

Cognitive Defusion Exercise

The defusion exercise used in the current study followed the procedures outlined by Masuda et al. (2009; 2010), and included a defusion rationale, training in defusion, and a 30 second repetition of a target word that was tracked by a stopwatch. Hayes et al. (2012) describe the rationale and training for the word repetition technique, which Masuda et al. (2009) conducted empirical research on and found that repeating a target word for at least 30 seconds led to decreases in emotional discomfort and believability of a problematic thought. The defusion rationale was presented to each participant with a brief description on the difference of fusion and defusion, after which participants were instructed on how to complete the WRT. Participants in both the HD and FD conditions were asked to say the word "milk" aloud and to notice all of the qualities of milk that came to mind upon saying the word, as well as anything else that came to mind when they said the word aloud. After discussing these qualities and how they represent fusion, participants were asked to repeat the world milk for 30 seconds and asked to notice what happened to their perceptions of the word as it was repeated. During the exercise the researcher

prompted the participants to adhere to the exercise by asking each participant after 10 seconds to repeat the word faster, and after 20 seconds prompting participants to repeat the word more loudly. Participants in the FD condition received an identical set of procedures focused on the word "lemon" following completing of the WRT for "milk." Participants in the CD condition followed the same procedures as the FD condition except that participants in this condition first completed the rationale and WRT for "car" before completing the WRT for "rabbit."

Statistical Analyses

Sample Size

The IRAP has been shown to successfully detect a small effect size with a relatively low number of participants (e.g., Hooper et al., 2010). Yet, in clinical research small effect sizes are not always meaningful. Thus, to determine an appropriate sample size for the current study a power analysis was conducted using G*Power Version 3.1 (Faul, Erdfelder, Buchner, & Lang, 2009; Faul, Erdfelder, Lang, & Buchner, 2007) based on the findings of Kishita et al. (2013) and Hooper et al. (2010) as they represent two studies assessing similar content as the current study. For Kishita and colleagues the study demonstrated an effect size of $f^2 = .301$, while Hooper and colleagues demonstrated an effect size of $f^2 = .262$. Using the parameters of a .261 effect size, an alpha level of .05, an 80% power level, and 4 conditions for a global effects MANOVA, G*Power predicted that a total sample of 48 (12 per condition) was needed to detect a similar effect in the current study. Such a small sample size was not sufficient for the purposes of the current study; therefore, the current study attempted to gather at least 120 participants (30 per condition), which would have been sufficient to capture an effect size of $f^2 = .150$.

Preliminary Analyses

Prior to completing any additional analyses, all collected data was reviewed and cleaned to maintain the integrity of the data. Data cleaning included checking the accuracy of data transfers into a single data set, calculating descriptive statistics and analyzing potential outliers, performing required transformations, and assessing for multicollinearity within measures (Tabachnick & Fidell, 2007). All self-report measures were compared between conditions to assess for comparability between conditions due to random assignment to condition.

Primary Analyses

Hypothesis 1: The IRAP will demonstrate sensitivity to changes in implicit cognition following a defusion intervention. Based on previous research showing that the IRAP can detect changes in implicit cognition following a therapeutic intervention (Hooper et al., 2010; Kishita et al., 2014; Ritzert et al., 2015), it was hypothesized that the IRAP would have:

- Significant differences in *D*-IRAP scores for trials containing the word "milk" between the control and half defusion condition
- Significant differences in *D*-IRAP scores for trials containing the word "milk" and "lemon" between the control math condition and full defusion condition
- Significant differences in *D*-IRAP scores between the control defusion and full defusion conditions

In order to test these predictions, a MANOVA was conducted to compare *D*-IRAP scores between the control math, half defusion, and full defusion conditions by using the four trial-type D-scores as dependent variables. These analyses enabled the researcher to determine if the defusion intervention impacted *D*-IRAP scores for the targeted stimuli of the intervention. Post hoc MANCOVAs were conducted as needed using the Food Attitudes Questionnaire, CFQ, DASS, and MPFI as covariates.

Hypothesis 2: IRAP response latencies will decrease on trials where the presented stimuli has undergone a defusion intervention. Based on research showing that a defusion intervention impacts IRAP performance by reducing response latencies for both consistent and inconsistent block-types (e.g., Kishita et al., 2014), it was hypothesized that:

- Significant differences in average response latency for trial-types of both block-types containing the word "milk" between the control math and half defusion conditions
- Significant differences in average response latency for all trial-types of both block-types between the control math and full defusion conditions
- Significant differences in average response latency for trial-types in both block-types between the control defusion and full defusion conditions

To test these predictions, a MANOVA was conducted to compare average response latencies between the control math, half defusion, and full defusion conditions for the four trial-types. The results of these tests allowed the researcher to determine if the average response latency on trials containing a word that has undergone defusion was significantly different than comparable trials in the control math condition that did not undergo a defusion intervention. Furthermore, the researcher was able to compare the defusion conditions to determine if training defusion for both milk and lemon (full defusion) produces statistically significantly reduced average latencies for all trials in comparison to a condition where defusion should theoretically impact only half of the trials (half defusion), or impact words not measured by the IRAP (control defusion). Post hoc MANCOVAs were conducted as needed using the Food Attitudes Questionnaire, CFQ, DASS, and MPFI as covariates.

CHAPTER 3

RESULTS

PRELIMINARY DATA CLEANING AND ANALYSES

After all data was collected a total of 122 participants had completed the study, and data compilation and cleaning began (identifying potential missing values, removal of outliers). Data was collected according to the methodology detailed in Chapter 2 above, and explicated in Appendix H. Once data collection finished, all questionnaire data was exported from Qualtrics to Excel where it was initially cleaned and organized. Additionally, all IRAP data (*D*-scores and latencies) were compiled into separate Excel files where they underwent viability checks to ensure all data had been appropriately collected and cleaned. Next, all Qualtrics and IRAP data were combined into a single SPSS file where multiple viability checks were completed to confirm that the data had been successfully merged. Thus, preliminary analyses were conducted on the full data set of 122 participants to assess for multicollinearity between measures and comparability between conditions on self-report measures of psychological functioning. When conducting primary and secondary analyses participants were excluded if they failed to meet 75% accuracy across all blocks of the IRAP, which left a total of 111 participants to be analyzed.

Demographic Information. To determine if conditions were comparable on demographics and attitudes and behavior toward milk and lemons, one-way ANOVAs were conducted. Results showed no statistical differences between groups on any demographic (all *p*s $\geq .108$). Of note, despite having three participants in both the CM and CD conditions with lactose intolerance there was no significant differences between conditions, $F(3, 118) = 2.073$, $p = .108$. Thus, preliminary analyses suggest that conditions were not significantly different on

demographics and attitudes/consumption of milk and lemon, and subsequent analyses do not include any of these items as potential covariates.

Examination of Self-Report Questionnaires

Attitudes Toward Milk and Lemon. All participants were asked to complete ratings of how often they consume milk and lemon on a weekly basis, as well as how much they enjoy milk and lemons. Mean scores were calculated for each condition (see Table 1), and a one-way ANOVA was conducted to compare conditions on consumption and attitudes for milk and lemon. Analyses indicated no significant differences between conditions for consumption of milk ($F(3, 118) = .627$, $p = .599$), or lemons ($F(3, 118) = .377$, $p = .769$). Similarly, there were no statistically significant differences between conditions on attitudes toward milk ($F(3, 118) = 1.045$, $p = .375$) or lemon ($F(3, 118) = 1.581$, $p = .198$). In addition, one-sample t-tests were conducted to determine if mean scores for these items were significantly different from zero, and analyses indicate that attitudes toward milk and lemon were both in the positive direction and significantly different from zero (both ps $< .001$), and that the number of days milk and lemon were consumed by participants in the sample were significantly greater than zero (both ps $< .001$). Taken together, these analyses indicate that no condition had significantly adverse attitudes or behaviors related to either milk or lemons that could be expected to impact IRAP performance.

Measures of Psychological Functioning. To assess whether conditions varied on different measures of psychological functioning, one-way ANOVAs were conducted for the CFQ, DASS 21 subscales, and MPFI subscales. Analyses showed that there were no significant differences between conditions on the CFQ, $F(3, 118) = 1.338$, $p = .265$. Likewise, there were no significant differences between conditions on the MPFI inflexibility subscale, $F(3, 118) = 1.644$,

p = .183, or MPFI flexibility subscale, $F(3, 118) = 1.048$, p = .374. However, analyses on the three subscales of the DASS 21 indicated statistically significant differences between groups on the Anxiety subscale, $F(3, 118) = 3.849$, p = .011, and Stress subscale, $F(3, 118) = 3.743$, p = .013, while the Depression subscale showed no significant differences, $F(3, 118) = 1.706$, p = .170. Post-hoc analyses for the Anxiety and Stress subscales showed that the CM condition was significantly different than the HD condition for both subscales, with a mean difference between conditions of 11.889 for Anxiety and 11.092 for Stress, indicating that CM participants had significantly higher levels of anxiety and stress than the HD participants on both subscales. Thus, participants across conditions were equivalent on the CFQ, MPFI, and DASS Depression subscale, but due to significant differences on the Anxiety and Stress subscales of the DASS 21 these subscales were considered as possible covariates in primary analyses. Moreover, a correlational analysis was conducted between all measures of psychological functioning to assess for collinearity (see Table 2), and due to the high level of correlation between DASS Anxiety and Stress subscales that is suggestive of collinearity, only DASS Anxiety was included in subsequent analyses.

IRAP Internal Consistency. In order to determine the internal consistency of the IRAP, a modified D_{IRAP} algorithm (see Appendix H) was used to produce two D-scores for the IRAP- one for odd-numbered trials and one for even-numbered trials. Pearson correlations were then conducted for the pair of D-scores for each condition, and a Pearson-Brown correction was calculated for each correlation to determine what internal consistency would be if IRAP trials were doubled. Results indicated that there was a significant correlation for the control math condition, $r(27) = .550$, p = .002, control defusion condition, $r(24) = .401$, p = .042, and half defusion condition, $r(25) = .721$, $p \leq .001$. These conditions demonstrated Spearman-Brown

corrections of $r = .71$, $r = .57$, and $r = .84$, respectively. However, there was no significant correlation for the full defusion condition, $r(27) = .258$, $p = .177$, with a Spearman-Brown correction of $r = .41$. These results show that only one of the four conditions (half defusion) demonstrates an internal consistency comparable to those reported by other IRAP studies (e.g., see Golijani-Moghaddam et al., 2013) before the Spearman-Brown correction, and after the correction all conditions but the FD condition demonstrates sufficient internal consistency.

Primary Analyses

Hypothesis 1: The IRAP will demonstrate sensitivity to changes in implicit cognition following a defusion intervention. To assess for changes in implicit cognition in response to a defusion intervention, a modified D_{IRAP} algorithm (see Appendix H) was used to produce IRAP D-scores for the four trial-types within each condition (see Table 3 for group means). These D-scores were then included in separate MANCOVAs, with DASS Anxiety as a covariate, to determine if the pattern of scores varied between conditions. To determine if the WRT for the word "milk" produced significant changes in D-scores, the first MANCOVA compared the control math condition (n = 29) to the half defusion condition (n = 27) for D-scores based on trials that include "milk" as the sample stimulus. Results show that there was no statistically significant difference in the pattern of D-scores between the conditions, $F(2, 54) = 2.116$, $p = .130$. The second MANCOVA compared the control math condition to the full defusion condition (n = 29) to assess whether completing the WRT for "milk" and "lemon" produced differences in the pattern of D-scores for all trial-types based on increased exposure to a defusion intervention. Analyses indicate that there were no significant differences in the pattern of scores between these conditions, $F(4, 54) = 1.724$, $p = .158$.

Following the MANCOVAs, secondary analyses were conducted to assess for significant differences between conditions on any D-score. First, an ANOVA was conducted to determine if the full defusion condition had significantly different D-scores in comparison to the control defusion condition (n = 26) for participants that completed the WRT for words not included in the IRAP. Analyses showed that there were no significant differences between conditions for the overall D-score, $F(1, 54) = .550$, $p = .462$, or any of the trial-type D-scores (all ps $\geq .064$). Next, an ANOVA was conducted to compare the half defusion condition to the full defusion condition to determine if completing the WRT for both sample stimuli produced significant differences, and analyses showed there was no statistical difference between conditions on any D-score (all ps $\geq .357$). Thus, based on the MANCOVAs and subsequent secondary analyses between conditions for all D-scores it appears that Hypothesis 1 in the current study was not supported, and results deviate from those reported in other IRAP studies (e.g., Hooper et al., 2010; Kishita et al., 2014).

Hypothesis 2: IRAP response latencies will decrease on trials where the presented stimuli has undergone a defusion intervention. To assess the effects of a defusion intervention on the pattern of response times in the IRAP, multiple MANCOVAs were conducted comparing the CM condition to the FD condition and the HD condition to the FD condition. As above, the DASS Anxiety subscale was included as a covariate in all analyses. In the first HD-FD comparison MANCOVA, only response latencies based on trials that had the word "lemon" as a sample stimulus were included to determine if the FD condition would have significantly different scores due to completing the WRT for the word "lemon." The second HD-FD comparison MANCOVA used response latencies for trials that had the sample stimulus "milk" to see if any significant differences existed between conditions that both completed the WRT for

"milk." Lastly, in the CM-FD comparison MANCOVA, all response latencies across all trials and blocks were utilized to assess for changes in the patterns of response latencies between the condition that received no defusion intervention and the condition that completed a defusion intervention for all IRAP sample stimuli (see Table 4 for response latency means for conditions).

Results showed that the MANCOVA comparing HD and FD conditions for trials with the sample stimulus "lemon" were non-significant, $F(4, 50) = .664$, $p = .620$. Similarly, the MANCOVA comparing these conditions for trials with the sample stimulus "milk" was non-significant, $F(4, 50) = .453$, $p = .770$. However, there was a significant difference present in the MANCOVA comparing the CM and FD conditions on the pattern of all response latencies, $F(8, 47) = 2.292$, $p = .037$. Based on this significant finding, follow-up one-way ANOVAs were conducted to determine which latency score(s) the CM and FD conditions differed on. Results of these analyses showed that there was a statistically significant difference between conditions on the response latency for trials with "milk" as the sample stimulus and pictures of milk as the target stimulus during consistent block-types, $F(1, 55) = 4.032$, $p = .050$. All other response latencies demonstrated no statistically significant difference between conditions (all ps $\geq .115$). Thus, the results indicate that the pattern of response latencies were significantly different between conditions, but that a single response latency (out of 8 possible) was driving this difference. Furthermore, these analyses fail to support Hypothesis 2 in the current study, deviating from previous IRAP research (Kishita et al., 2014).

Exploratory Analyses

Milk/Lemon Defusion Measure. To determine if conditions significantly differed post-experiment in response to the words "milk" and "lemon" participants completed items rating the saliency of fused and defused responses for "milk" and "lemon," and ANOVAs were conducted

to compare conditions. Results indicated that there were no statistical differences between conditions on responses that represented fusion (p = .366) or defusion (p = .654) for "milk." Similarly, there were no significant differences between conditions for fused (p = .590) or defused (p = .668) responses for "lemon." Thus, evidence suggests that the defusion exercises for the CM, HD, or FD conditions did not impact self-reported defused responses for stimuli in the current study. Interestingly, group means for each of the fused and defused responses (see Table 5) do show that the control math condition, in comparison to defusion conditions, had higher scores indicative of fusion except for the "lemon" vs. "yellow" comparison. While these group means do not indicate statistical significance of any kind, it may provide some evidence that the defusion conditions were responding to fused reactions, but not defused responses, differently than the control math condition.

Multidimensional Psychological Flexibility Inventory. Despite the MPFI flexibility and inflexibility scales showing no significant differences between conditions, follow-up analyses were conducted to see if specific subscales of the MPFI could account for unexplained variance between conditions. MANCOVAs were conducted on latency scores and D-scores with the defusion and fusion subscales of the MPFI as covariates. Analyses showed that there were no significant differences between conditions for response latencies with the defusion subscale as a covariate, $F(24, 300) = 1.190$, p = .249, or with the fusion subscale as a covariate, $F(24, 300) = 1.171$, p = .267. Moreover, there were no significant differences between conditions on D-scores with the defusion subscale as a covariate, $F(12, 315) = .994$, p = .455, or the fusion subscale, $F(12, 315) = .979$, p = .469.

CHAPTER 4

DISCUSSION

Problematic cognitions are a primary factor in numerous psychological disorders. Due to their influence in so many disorders, problematic cognitions have been a focus of clinical research and interventions over the last 50 years, and several therapies have devised approaches to target problematic cognitions. As stated in Chapter 1, Cognitive Behavior Therapy has been the most widely studied and supported therapy to treat psychological disorders and distorted cognitions through cognitive restructuring (e.g., Bryant et al., 2003; Mattick & Peters, 1988). Yet, other therapeutic techniques have provided evidence that cognitive restructuring and similar cognitive therapy approaches are not the only method of alleviating the impact of problematic cognitions on psychological health. For example, cognitive defusion from Acceptance and Commitment Therapy assists individuals in reducing the emotional impact of thoughts by deliteralizing the content and meaning of thoughts so that one can have a thought without rigidly acting on (or in accordance with) the thought (Arch et al., 2012; Blackledge & Drake, 2013; Hayes, Strosahl, & Wilson, 2012). Over the last couple of decades researchers have found the cognitive defusion is an effective (Masuda et al., 2004; 2010) and comparable (Arch et al., 2012) alternative to cognitive restructuring in the treatment of problematic cognitions.

To assess the effectiveness of different therapeutic interventions, such as cognitive restructuring or defusion, researchers and clinicians employ self-report measures of psychological functioning to assess changes from pre- to post-intervention. While this method has been utilized for decades to determine the success of psychological treatments, recent research on self-report measures have found multiple issues of concern as they relate to the assessment of cognitive processes. For example, several researchers have captured what is

referred to as the repeated measures effect (Choquette & Hesselbrock, 1987; Gilbert et al., 2002), while others have identified and discussed at length issues around the social desirability effect and response bias (e.g., Fleming, 2012; Hanley, 1961; Kuentzel, Henderson, & Melville, 2008). When taken together, these different issues pose a threat to the reliability and validity of self-report measures of psychological functioning. In contrast to self-report measures, several behavioral measures of cognitive processes (i.e., implicit measures) have been developed to detect changes in psychological functioning due to therapeutic interventions (e.g., Dasgupta & Greenwald, 2001; Hooper et al., 2010; Kishita et al., 2014). Similar to self-report measures, implicit measures are often administered pre- and post-intervention to assess for potential changes in psychological functioning or distress, and studies have shown the IRAP to be sufficiently sensitive to detect changes in implicit attitudes and psychological functioning (e.g., Hooper et al., 2010; Kishita et al., 2014). Moreover, the benefit of utilizing implicit measures of cognition is that they assume that implicit cognitions are difficult to fake on reaction-time tasks because such attitudes are outside of conscious awareness and can capture a unique source of information about psychological functioning that cannot be measured by self-reports alone.

While incorporating the research on cognitive defusion and the IRAP, the current study attempted to expand the research findings on the IRAP by showing that the IRAP could detect differences in implicit cognition following the implementation of the WRT (Titchener, 1916) for stimuli used in the IRAP. First, this study endeavored to replicate previous findings showing that the IRAP is capable of measuring differences in implicit cognitions (i.e., D-scores) following a defusion intervention (e.g., Hooper et al., 2010; Kishita et al., 2014), and that the defusion intervention would affect IRAP D-scores in a hypothesized pattern. Second, the current study attempted to contribute to the literature showing that defusion interventions alter reaction times

for trial-types in the IRAP (e.g., Kishita et al., 2014) by reducing response latencies for both consistent and inconsistent block-types. Below is a description of key findings and implications of the current study, and a discussion of the study limitations, strengths, and potential directions of future research on the IRAP are included.

Summary and Discussion of Findings

Hypothesis One

A growing body of research on defusion has shown that defusion interventions decrease perceived levels of distress and believability toward problematic cognitions by deliteralizing (Blackledge & Drake, 2013) the meaning of a word or thought from the sensory details of the word(s). Studies demonstrating such effects commonly use interventions that are conducted in psychotherapy without controlling for the quality or duration of the intervention. Thus, the current study provides an ideal opportunity to more strongly control how a defusion intervention is implemented and for how long, while also offering an opportunity to assess the effectiveness of a highly controlled intervention on IRAP performance. To test the hypothesis that attitudes toward "milk" and "lemon" would be impacted by completing the WRT (i.e., "milk exercise") as a defusion intervention, participants were randomly assigned to four conditions- three of which required participants to complete the WRT for (1) words unrelated to stimuli in the IRAP, (2) for only the word "milk", or (3) for the words "milk" and "lemon." Following completion of the intervention phase of the study (WRT or simple math problems), participants were administered a standard IRAP with the words "milk" and "lemon" as sample stimuli and various pictures of milk and lemons as target stimuli. Based on this design, the focus of the study was to determine if a defusion intervention significantly impacted IRAP performance in comparison to a control condition that did not received the WRT. Additionally, by including three defusion conditions

the current study also allowed for comparisons between conditions on the impact of the WRT for different stimuli and dosages on a standard IRAP.

Results of various MANCOVAs revealed no significant differences in the pattern of IRAP D-scores between any of the four conditions. This finding holds true for analyses comparing the pattern of D-scores between the CM condition and the HD condition, as well as between the CM and FD conditions. Moreover, there were no significant differences between any of the defusion conditions, indicating that completing the WRT for stimuli irrelevant to the IRAP (i.e., CD condition) or completing the WRT for only one of the IRAP stimuli (i.e., HD condition) did not have a significant impact on IRAP performance in comparison to the FD condition. To be thorough, follow-up ANCOVAs were completed to determine if conditions significantly differed on any single IRAP D-score, instead of the pattern of all D-scores, and results showed no significant differences between conditions on the four trial-types or overall D-score. Therefore, the current study failed to detect any differences in implicit attitudes toward milk and lemon following defusion interventions focused on "milk" and "lemon" across any of the four conditions.

These findings deviate from those of previous IRAP studies demonstrating the IRAP's success in measuring changes in implicit cognition following a defusion intervention (e.g., Hooper et al., 2010; Kishita et al., 2014). One of the defining premises of the current study was that simple and highly familiar words that have a long and repetitive learning history would exemplify fusion, consistent with RFT, and would be prime candidates for a defusion intervention focused on altering responses toward strongly fused words. However, when considering the current study in comparison to other IRAP studies it seems plausible that defusion interventions did not impact IRAP D-scores because milk and lemon are not

emotionally evocative stimuli, and defusion is posited to reduce the emotional impact and believability of a word or thought (e.g., Masuda et al., 2009). In other words, completing a defusion intervention such as the WRT on words that do not elicit strong emotional reactions may not (or at least not as quickly) demonstrate a defusion effect as measured by a reaction-time task. Furthermore, participants who completed the WRT did so only once for the targeted word(s) within a three to four-minute period, and this may not have been a sufficient dose of defusion to produce significant differences in D-scores. Thus, one might expect a defusion intervention to have stronger effects on IRAP performance if the IRAP was measuring implicit attitudes toward stimuli with stronger emotional valences than "milk" and "lemon," and if participants completed multiple rounds of defusion for the stimuli being measured by the IRAP. Moreover, IRAP D-scores may not represent the most effective way to assess for defusion effects in the IRAP due to the methodology used to calculate D-scores (see Appendix H). The standardization procedure for calculating D-scores provides a limited way to assess IRAP performance, and other approaches to analyzing IRAP data could provide broader information on possible defusion effects.

Hypothesis Two

Of the few IRAP studies that have utilized defusion interventions to alter implicit cognition, only Kishita et al. (2014) reported the impact that a defusion intervention demonstrated on response latencies. Based on the findings of Kishita and colleagues (2014) showing that defusion reduced response latencies for both consistent and inconsistent block-types, the current study sought to replicate these findings and demonstrate that the WRT would significantly impact response latencies for trials that included the word(s) targeted by the WRT across the various defusion conditions. Specifically, this study hoped to show that: (1) the

defusion conditions would demonstrate significantly faster response times on appropriate trials when compared to the CM condition, (2) receiving defusion for one sample stimulus (HD condition) would reduce latencies for only trials that included that stimulus when compared to the FD condition, and (3) the FD condition would demonstrate significantly reduced response latencies when compared to a condition that received the WRT for an equivalent duration but for words not included in the IRAP (CD condition).

Results of the MANCOVAs testing this hypothesis showed no significant differences between the HD and FD conditions on the pattern of response latencies for either "milk" or "lemon" trials, and follow-up analyses confirmed that there were no differences between the HD and FD conditions on any of the response latencies for trail-types or block-types. However, there was a significant difference in the pattern of response latencies between the CM and FD conditions. Follow-up one-way ANCOVAs revealed a significant difference on trials showing the word "milk" with pictures of milk during consistent block-types (i.e., requiring participants to say "milk" is the "same" as pictures of milk), but no other trial-type displayed significant differences regardless of block-type. These findings show that the findings of the MANCOVA were being driven by a single significant difference among all response latencies included in the analysis, which appeared to be artificially inflating the significance of the MANCOVA finding.

Similar to hypothesis one, the findings for the current hypothesis are inconsistent with those reported by Kishita et al. (2014). If one considers the possible explanations proposed above regarding hypothesis one's non-significant findings, it makes sense why response latencies appear comparable across the conditions for the various trial- and block-types. This is to say that if emotionally evocative words or thoughts are necessary to demonstrate a defusion effect in the IRAP, then it can be expected that response latencies will be unaffected by a defusion

intervention if the stimuli do not follow this supposition. Additionally, completing only one IRAP likely limited the effect of defusion interventions on response latencies. If participants were to complete multiple defusion interventions over time one could expect to see greater differences in latencies if a true defusion effect were present. Furthermore, completing multiple IRAPs over time would elucidate if a defusion effect is delayed in impacting IRAP performance.

Other possible explanations for these non-significant findings can also be considered. First, it may make sense why the milk-milk trial-type in consistent blocks was most strongly affected by a defusion intervention when compared to other trial-types. Given the two stimuli included in the IRAP – milk and lemon – individuals may have had many more learning experiences with milk in their lives than with lemons, providing more opportunities for a fused response to develop to milk-related stimuli and provide a stronger opportunity for a defusion effect to occur from the WRT. This can be somewhat corroborated by the Food Attitudes Questionnaire included in the current study which showed that milk had been consumed an average of 2.2 days in the week prior to the current study across all participants, whereas lemon had been consumed on average less than one day in the previous week. If extrapolated to include all possible situations in which individuals are exposed to milk or lemons throughout life, exposure to milk appears likely to occur thousands more times than exposure to lemon which would theoretically bolster the fusion individuals have with milk stimuli. Thus, one might expect a stronger baseline fusion response to stimuli would more easily be affected by a defusion intervention. Second, given the magnitude of non-significant findings for response latencies, especially in comparison to Kishita et al. (2014), it is possible that the single significant finding reported above is due to Type 1 error (Tabachnick & Fidell, 2013). This explanation is more

reasonable than others given that in all analyses conducted for the purposes of the current study, only one significant finding of any kind was identified.

Overall, the current findings fail to provide evidence that the IRAP is sensitive enough to measure differences in implicit cognition due to a defusion intervention. Furthermore, the current study also provides minimal evidence that a defusion intervention can significantly affect response latencies as measured by the IRAP. Despite a lack of statistical support for the a priori hypotheses in the current study, comparing the average response latency for each trial-type and block-type from Table 4 shows that the FD and CD conditions have lower response latencies than the CM condition on all latencies except for one trial-type comparison. Although the differences between average latencies is not statistically significant, it does indicate a slight trend that is consistent with the findings of Kishita et al. (2014) and suggests the defusion intervention had a small effect on response latencies and could clarify a direction for future research on using defusion interventions with the IRAP. A final explanation for the current findings points to a larger issue identified in this study- IRAP response latencies may not exemplify defusion as strongly as suggested in previous research. One reason for this is that implicit attitudes might not reflect fusion or defusion at all, which would indicate that the IRAP can be considered a simple sorting task that individuals get better at with practice and has no bearing on overt behavior. A second reason would be that latencies reflecting defusion (or fusion) do not translate into commensurate behaviors outside the IRAP. For example, responding more quickly on IRAP trials could be considered a defused response (e.g., Kashita et al., 2013) but not predict the likelihood that an individual would engage in behavior resembling defusion in real life. If either of these explanations are true they would invalidate the use of the IRAP as a metric for defusion, and possibly other ACT processes.

Strengths, Limitations, and Implications for Future Research

The current study contributes to the literature on the effects of defusion on the IRAP, but this study also has several strengths and weaknesses that should be considered. Some characteristics of the research design offered advancements over previous IRAP and defusion research. For instance, the current study displays a strong connection to RFT principles regarding defusion (see Chapter 1 above), and utilized common stimuli that were expected to capture fused and defused responses in accordance with the learning principles discussed in RFT. .In addition, this study attempted to exhibit strong control over the administration and dosage of a defusion intervention by varying the administration of the WRT to assess its effects on IRAP performance. Moreover, the current study also helps determine the impact of completing the WRT for words not assessed in the IRAP by ascertaining if simple exposure to defusion can account for differences in IRAP performance. Lastly, this study incorporated self-report measures with the IRAP to assess for differences between conditions on overt and implicit attitudes, providing an opportunity to capture unexplained variance unaccounted for by either method individually. These factors, when taken together, provide a strong foundation for the design and implementation of the current study to assess meaningful, theoretically-based questions that pertain to the IRAP and cognitive defusion.

As mentioned above, few studies have discussed the effects of psychological interventions on block-type response latencies in the IRAP, and this study contributes to that body of literature. Kishita, Ohtsuki, Sakai, and Muto (2010) provided what appears to be the first study looking at the effects of a defusion intervention on IRAP latencies and found that defusion reduced response latencies post-intervention for both block-types, with the greatest reductions occurring for inconsistent blocks. Kishita et al. (2014) replicated this finding for both block-

types of the IRAP. Drake et al. (2016) who found a similar finding in that individuals who reported higher levels of defusion (but who did not complete a defusion intervention) responded faster on both block-types of the IRAP than those reporting lower levels of defusion (i.e., higher fusion). The current study adds to this group of studies, albeit with less power and statistical significance, by showing that a defusion intervention not only decreased block-type response latency, but also decreased response latencies for all trial-types across both block-types in comparison to a control condition. Future researchers could follow this trend and more closely analyze the effects of a defusion intervention on both IRAP block-type latencies and individual trial-types within block-types. Such an approach would provide a more refined understanding of the impact of defusion on implicit attitudes and enhance the field's understanding of the clinical impact of defusion interventions.

The current study also had several limitations worth discussing. First, although the current study had more participants per condition than other IRAP studies (e.g., Hooper et al., 2010; Kishita et al., 2013; Nicholson & Barnes-Holmes, 2012), the total sample size was not sufficiently large enough to capture a defusion effect in IRAP performance. A total of 122 participants completed the study, with an average of 30 participants per condition, which met preliminary G-power analyses suggesting 30 participants per condition were required to obtain appropriate power and medium effect sizes if present. Post-hoc analyses to determine what effect size could be appropriately captured by the current study found that the sample size was large enough to measure an $f^2 = .100$, but the actual post-hoc effect size of this study was .035. Such a finding indicates that if a true defusion effect existed the current design and number of participants should have been able to capture it in some capacity, but the sample size failed to capture an appropriate effect size. However, a larger sample size (e.g., 50 to 75 participants per

condition) would provide higher statistical power for the study, and a more conclusive analysis on if the IRAP can accurately determine how a defusion effect would impact IRAP D-scores and response latencies.

An additional issue to consider is that previous IRAP research assessing the effects of a defusion intervention on implicit cognition assess such effects with a repeated-measures design, while the current study utilized a between-subjects design. Thus, the current study can only draw conclusions about differences between conditions following an intervention and cannot conclude if the defusion exercise had a significant effect on IRAP performance. This decision was made to shorten the time necessary to complete the study, thus enabling a full sample to be obtained within the time constraints for this study. However, this limits the comparison that can be made between the current results and the results reported by previous IRAP studies (e.g., Kishita et al., 2014; Ritzert et al., 2014). Future research would benefit from a repeated-measures design implementing the same procedures as the current study as it would allow for an additional level of analyses (i.e., within-subjects) assessing the effectiveness of a defusion intervention on IRAP performance.

Another limitation of this study is that the defusion intervention focused on altering attitudes toward innocuous stimuli (i.e., milk and lemon) whereas clinical applications of defusion address evaluative/judgmental content of an individual's thoughts (e.g., "I am WORTHLESS," "Everyone is DANGEROUS"). The decision to defuse from "milk" and "lemon" was based on assumption consistent with RFT that well learned repertoires of responding (e.g., saying or thinking "milk" every time someone sees milk throughout one's life) are prime opportunities to apply defusion and loosen the derived relational responding. However, because the content of the IRAP stimuli are not clinically relevant and do not elicit strong

emotional reactions, the use of the WRT in the current study likely did not provide a defusion experience comparable to psychotherapy practices, thus limiting the defusion effect present in the current data in comparison to other IRAP studies (e.g., Kishita et al., 2014). A future study could implement the same design as the current study and incorporate stimuli into the IRAP that represent clinically relevant problematic cognitions (e.g., thoughts common with anxiety or depression), and would be ideal targets for a defusion intervention consistent with those addressed in psychotherapy. Such a study would enable researchers to more thoroughly assess if IRAP effects are enhanced by the inclusion of clinically relevant stimuli, and how a clinically appropriate defusion intervention impacts IRAP D-scores and response latencies. These findings could help elucidate how defusion leads to behavior change, thus offering opportunities to strengthen defusion interventions in treatment. However, it should be noted that if the IRAP requires emotionally evocative stimuli in order to obtain a defusion effect it would indicate that the IRAP is not an appropriate measurement for RFT and ACT principles as defusion is presumed to function on all verbal stimuli regardless of emotional relevance. Therefore, if future studies found results similar to the current study it would suggest either: a) RFT and assumptions about how defusion works on cognition is wrong and needs to be reassessed, or b) the IRAP is an inconsistent behavioral measure of implicit cognition and should not be used to assess the effects of psychological interventions.

An additional weakness is that there is no way to assure that participants were adequately engaging in the WRT. Due to the nature of the stimuli used in this study, and the limited time allotted to complete the WRT with a rationale and experiential exercise, it's feasible that participants complete the intervention as instructed without truly experiencing defusion. One way to limit this possibility would be to conduct an idiographic study in which stimuli used in

the study are selected by participants for their emotional relevance, thus allowing researchers to better assess if defusion is occurring during the WRT (e.g., Masuda et al., 2004; 2008). This method was not implemented because of constraints on time and resources but would create a more powerful study. Another possible concern with the WRT is whether all experimenters were adequately trained in the implementation of the intervention. Researchers conducted a training during the planning of the study to ensure proper procedures were disseminated to all experimenters regarding the WRT, and guides were given to all experimenters to follow and track completion of each component of the study. However, there is no way to determine if there was strict adherence to study procedures for the WRT across all experimenters. Subsequent studies could address this in two ways: (1) by employing a more rigorous training period in which all experimenters are assessed on their administration of the study protocol until meeting a predetermined standard, at which point experimenters could begin live administrations with participants; and (2) video recording the administration of the WRT so that all participants receive a standardized administration of the intervention to eliminate the possibility of between-experimenter effects for the WRT. Having said all of this, it should be noted the current study did not demonstrate any significant differences between experimenters for any component (self-reports, IRAP, post-intervention measures) of the study, but the methodology could be enhanced in future research.

Furthermore, the current study utilizes only undergraduate students at a mid-western American university, made up of primarily Caucasian, 18 to 19-year-old students who have not engaged in therapy previously. Due to these demographics the relevance of this study is restricted to similar homogenous undergraduate samples, and it lacks generalizability to a clinical sample of individuals who would most benefit from research on the effects of cognitive

defusion. Therefore, if future researchers intended to expand the literature on defusion as it relates to the IRAP, the use of a clinical sample would offer ample opportunity to increase the external validity of this research by measuring (hopefully) changes in performance in an IRAP representative of one's psychological distress. This could be achieved by asking individuals during a waitlist period before beginning therapy to identify problematic thoughts they experience to be used as sample stimuli, and the emotional and cognitive reactions they experience in response to such thoughts as target stimuli. Then, these individuals could be guided through the WRT for the thoughts they struggle with before completing the IRAP to assess for changes in implicit cognition for these thoughts. Moreover, some participants could be asked to complete the WRT and IRAP multiple times to determine if there is a dosage effect necessary for defusion to be captured by IRAP performance. Results of this kind of study would greatly advance our understanding of how therapeutic interventions influence implicit cognition and behavioral responses to problematic thoughts.

After the WRT, participants engaged the IRAP and then completed self-reports intended to measure defusion effects. Neither type of measure revealed group differences, and so it appears that either the WRT did not generate defusion effects or the measures were not adequately sensitive to it. However, it seems conceivable that the self-reports did not detect any defusion effects because such effects wore off by the time the measure was administered. It may have been preferable to counterbalance the order of the IRAP and the self-reports post-intervention, at least to provide a means of checking for order effects. This option was not incorporated in the current study because of the emphasis on detecting group differences with the IRAP, but future studies might consider counterbalancing the order of measures or adding subsequent conditions that receive multiple defusion interventions and IRAPs followed by self-

report assessments of defusion and psychological flexibility following each component of the study.

Conclusion

In summary, the current study failed to provide evidence that a defusion intervention can account for significant changes in IRAP performance as measured by IRAP D-scores. However, this study did provide initial evidence, consistent with other IRAP studies, that a defusion intervention led to significant differences in the pattern of response latencies on the IRAP. These findings were clarified by post-hoc analyses showing that a single trial-type within consistent IRAP blocks was driving the difference in response latencies between conditions. This brings into question why only one response latency was statistically significant between conditions and suggests that any significant findings cannot solely be explained by the defusion intervention that was implemented, and indicates the presence of Type 1 error in the analyses. This study was designed and implemented in a manner consistent with a good understanding of Relational Frame Theory and cognitive defusion from Acceptance and Commitment Therapy. In addition, the current study design provides a strong example of how IRAP researchers can test the influence of therapeutic interventions on implicit attitudes, albeit without utilizing a clinical sample. Although the findings of this study failed to meet expectations, the literature on the IRAP and therapeutic interventions need studies such as this one to rigorously test the assumptions underlying psychological treatments and their impacts on human cognition and behavior.

Table 1

Means (Standard Deviation) for attitudes toward milk and lemon

Item	CM	CD	HD	FD
Days of consuming milk in last week	2.00 (2.04)	2.38 (2.21)	2.52 (2.14)	1.86 (2.07)
Days of consuming lemon in last week	1.03 (1.52)	1.08 (2.08)	0.48 (1.19)	0.90 (1.54)
Likability of milk	0.14 (1.85)	0.85 (1.69)	0.70 (1.94)	0.83 (1.69)
Likability of lemon	1.00 (0.14)	0.27 (2.16)	-0.11 (1.74)	0.72 (1.73)

Note. Items assessing likability are based on a -3 to 3 Likert scale.

Table 2

C*orrelations (N) between measures of psychological functioning*

Measure	1	2	3	4	5	6
1. CFQ	1	.66** (122)	.72** (122)	.67** (122)	.82** (122)	-.33** (121)
2. DASS-d		1	.89** (122)	.90** (122)	.71** (122)	-.29** (121)
3. DASS-a			1	.92** (122)	.75** (122)	-.34** (121)
4. DASS-s				1	.72** (122)	-.31** (121)
5. MPFI-i					1	-.36** (121)
6. MPFI-f						1

Note. CFQ = Cognitive Fusion Questionnaire; DASS-d = Depression, Anxiety and Stress Scale-depression subscale; DASS-a = Depression, Anxiety and Stress Scale-anxiety subscale; DASS-s = Depression, Anxiety and Stress Scale-stress subscale; MPFI-i = Multidimensional Psychological Flexibility Inventory-inflexibility subscale; MPFI-f = Multidimensional Psychological Flexibility Inventory-flexibility subscale. ** $p \leq .001$.

Table 3

Means (Standard Deviation) of Implicit Relational Assessment Procedure D-scores

Condition	CM	CD	HD	FD
IRAP D-score				
Milk-Milk	.40 (.33)	.42 (.31)	.52 (.40)	.56 (.25)
Milk-Lemon	.27 (.38)	.33 (.39)	.23 (.30)	.21 (.37)
Lemon-Lemon	.40 (.35)	.31 (.32)	.32 (.33)	.41 (.39)
Lemon-Milk	.33 (.33)	.22 (.41)	.19 (.41)	.26 (.42)
Overall	.35 (.23)	.32 (.21)	.31 (.26)	.36 (.23)

Note. CM = Control Math condition, CD = Control Defusion condition, HD = Half Defusion condition, FD = Full Defusion condition.

Table 4

Means (Standard Deviation) of Implicit Relational Assessment Procedure response latencies

Condition Block/trial-type	CM	CD	HD	FD
Consistent/Milk-Milk	1391.58 (304.52)	1300.70 (242.91)	1338.06 (283.02)	1243.32 (251.21)
Inconsistent/Milk-Milk	1643.59 (414.43)	1536.63 (256.86)	1627.33 (345.74)	1538.96 (421.66)
Consistent/Milk-Lemon	1507.61 (361.69)	1402.44 (265.26)	1557.79 (304.09)	1479.03 (340.59)
Inconsistent/Milk-Lemon	1627.02 (353.57)	1588.47 (292.46)	1714.75 (365.42)	1641.52 (416.25)
Consistent/Lemon-Lemon	1421.97 (333.35)	1348.05 (247.73)	1404.93 (325.12)	1302.67 (298.70)
Inconsistent/Lemon-Lemon	1684.27 (440.49)	1510.13 (285.81)	1591.40 (381.99)	1514.87 (355.88)
Consistent/Lemon-Milk	1498.03 (330.69)	1416.91 (252.20)	1519.83 (361.22)	1391.41 (264.16)
Inconsistent/Lemon-Milk	1704.81 (375.40)	1503.99 (255.34)	1642.45 (354.23)	1576.02 (428.58)
Overall Consistent	1454.79 (305.65)	1367.15 (217.73)	1455.15 (298.14)	1354.11 (264.85)
Overall Inconsistent	1664.92 (369.15)	1534.98 (246.49)	1643.98 (337.33)	1567.67 (357.50)

Note. CM = Control Math condition, CD = Control Defusion condition, HD = Half Defusion condition, FD = Full Defusion condition.

Table 5

Means (Standard Deviation) for Milk-Lemon defusion measure

Condition	CM	CD	HD	FD
Items				
Milk vs.				
Cow	8.72 (2.03)	8.42 (2.58)	8.00 (2.70)	7.90 (3.05)
Creamy	6.41 (3.09)	6.35 (3.17)	5.89 (3.22)	4.69 (3.38)
Dairy	8.59 (2.67)	8.38 (3.09)	7.78 (2.72)	7.38 (3.16)
Sound	2.52 (3.32)	2.31 (2.99)	1.96 (2.85)	3.34 (3.52)
Symbol	3.24 (3.55)	2.92 (3.32)	2.48 (3.06)	3.34 (3.58)
Text	3.86 (3.83)	2.38 (3.26)	2.56 (3.14)	3.69 (3.97)
White	8.72 (2.17)	8.69 (2.09)	8.07 (2.65)	7.79 (2.83)
Word	5.21 (3.70)	4.69 (3.82)	4.04 (3.67)	4.83 (3.37)
Lemon vs.				
Citrus	8.21 (2.87)	7.65 (3.56)	7.37 (3.19)	7.34 (3.62)
Fruit	7.69 (3.14)	7.00 (3.67)	7.26 (2.80)	7.52 (3.37)
Juicy	7.38 (3.25)	7.46 (3.35)	6.26 (2.65)	6.62 (3.13)
Sound	2.03 (3.08)	1.62 (2.45)	1.93 (2.51)	3.41 (3.40)
Symbol	3.07 (3.74)	3.35 (3.21)	2.19 (2.86)	3.62 (3.81)
Text	3.79 (3.87)	2.69 (3.51)	2.56 (3.03)	3.24 (3.51)
Word	5.00 (4.07)	4.19 (3.87)	4.00 (3.77)	4.17 (3.73)
Yellow	8.83 (2.24)	9.35 (1.164)	8.81 (2.35)	8.72 (2.59)
Sum Totals				
Milk defusion	15.10 (12.01)	13.42 (9.98)	12.07 (11.05)	15.10 (10.71)
Milk fusion	31.65 (7.45)	29.39 (10.69)	29.23 (8.28)	27.60 (8.83)
Lemon defusion	14.39 (13.07)	12.65 (13.34)	11.20 (10.57)	14.30 (12.03)
Lemon fusion	32.32 (9.17)	29.55 (11.28)	29.07 (9.42)	29.93 (10.08)

Note. CM = Control Math condition, CD = Control Defusion condition, HD = Half Defusion condition, FD = Full Defusion condition.

REFERENCES

American Psychiatric Association. (2013). Diagnostic and statistical manual of mental disorders (5th ed.). Arlington, VA: American Psychiatric Publishing.

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,* 469-478. doi:10.1016/j.brat.2012.04.007

Avdagic, E., Morrissey, S. A., & Boschen, M. J. (2014). A randomized controlled trial of acceptance and commitment therapy and cognitive-behaviour therapy for generalized anxiety disorder. *Behaviour Change, 31(2)*, 110-130. doi: 10.1017/bec.2014.5

Barnes-Holmes, D., Barnes-Holmes, Y., Power, P., Hayden, E., Milne, R., & Stewart, I. (2006). Do you really know what you believe? Developing the implicit relational assessment procedure (IRAP) as a direct measure of implicit beliefs. *The Irish Psychologist, 32(7),* 169–177.

Barnes-Holmes, D., Barnes-Holmes, Y., Stweart, I., & Boles, S. (2010). A sketch of the implicit relational assessment procedure (IRAP) and the relational elaboration and coherence (REC) model. *The Psychological Record, 60*, 527-542.

Barnes-Holmes, D., Murphy, A., Barnes-Holmes, Y., & Stewart, I. (2011). The implicit relational assessment procedure: Exploring the impact of private versus public contexts and the response latency criterion on pro-white and anti-black stereotyping among white Irish individuals. *Psychological Record, 60(1),* 57–79.

Beck, A. T. (1967). Depression: Causes and treatment. Philadelphia, PA: University of Pennsylvania Press.

Beck, J. S. (1995). Cognitive therapy: Basics and beyond. New York, NY: The Guilford Press.

Beck, J. S. (2011). Cognitive behavior therapy: Basics and beyond (2nd ed.). New York, NY: The Guilford Press.

Biran, M. & Wilson, G. T. (1981). Treatment of phobic disorders using cognitive and exposure methods: A self-efficacy analysis. *Journal of Consulting and Clinical Psychology, 49(6),* 886-899.

Birnbaum, H. G., Kessler, R. C., Kelley, D., Ben-Hamadi, R., Joish, V. N., & Greenberg, P. E. (2010). Employer burden of mild, moderate, and severe major depressive disorder: Mental health services utilization and costs, and work performance. *Depression and Anxiety, 27*, 78-89. doi: 10.1002/da.20580

Blackledge, J. T. (2007). Disrupting verbal processes: Cognitive defusion in acceptance and commitment therapy and other mindfulness-based psychotherapies. *The Psychological Record, 57,* 555-576.

Blackledge, J. T. (2015). Cognitive defusion in practice: A clinician's guide to assessing, observing & supporting change in your client. Oakland, CA: Context Press.

Blackledge, J. T., & Drake, C. E. (2013). Acceptance and commitment therapy: Empirical and theoretical considerations. In S. Dymond & B. Roche (Eds.), *Advances in Relational Frame Theory and Contextual Behavioral Science: Research and Application* (pp. 219-252). Oakland, CA: New Harbinger.

Blair, I. V., Ma, J. E., & Lenton, A. P. (2001). Imagining stereotypes away: The moderation of implicit stereotypes through mental imagery. *Journal of Personality and Social Psychology, 81(5),* 828-842. doi: I0.1037//0022-3514.81.5.828

Bluett, E. J., Homan, K. J., Morrison, K. L., Levin, M. E., & Twohig, M. P. (2014). Acceptance and commitment therapy for anxiety and OCD spectrum disorders: An empirical review. *Journal of Anxiety Disorders, 28*, 612-624. http://dx.doi.org/10.1016/j.janxdis.2014.06.008

Bruchmuller, K., Margraf, J., Suppiger, A., & Schneider, S. (2011). Popular or unpopular? Therapists' use of structured interviews and their estimation of patient acceptance. *Behavior Therapy, 42(4)*, 634-643. http://dx.doi.org.proxy.lib.siu.edu/10.1016/j.beth.2011.02.003

Bryant, R. A., Moulds, M. L., Guthrie, R. M., Dang, S. T., Mastrodomenico, J., Nixon, R. D. V., Felmingham, K. L., Hopwood, S., & Creamer, M. (2008). A randomized controlled trial of exposure therapy and cognitive restructuring for posttraumatic stress disorder. *Journal of Consulting and Clinical Psychology, 76(4),* 695-703. doi: 10.1037/a0012616

Bryant, R. A., Moulds, M. L., Guthrie, R. M., Dang, S. T., & Nixon, R. D. V. (2003). Imaginal exposure alone and imaginal exposure with cognitive restructuring in treatment of posttraumatic stress disorder. *Journal of Consulting and Clinical Psychology, 71(4),* 706-712. doi: 10.1037/0022-006X.71.4.706

Butler, A. C., Chapman, J. E., Forman, E. M., & Beck, A. T. (2006). The empirical status of cognitive-behavioral therapy: A review of meta-analyses. *Clinical Psychology Review, 26,* 17-31. doi:10.1016/j.cpr.2005.07.003

Carpenter, K. M., Martinez, D., Vadhan, N. P., Barnes-Holmes, D., & Nunes, E. V. (2012).

Measures of attentional bias and relational responding are associated with behavioral treatment outcome for cocaine dependence. *The American Journal of Drug and Alcohol Abuse, 38,* 146-154.

Choquette, K. A. & Hesselbrock, M. N. (1987). Effects of reteting with the Beck and Zung depression scales in alcoholics. *Alcohol & Alcoholism, 22(3),* 277-283.

Cullen, C., Barnes-Holmes, D., Barnes-Holmes, Y., & Stewart, I. (2009). The implicit relational assessment procedure (IRAP) and the malleability of ageist attitudes. *The Psychological Record, 59*, 591-620.

Dasgupta, N. & Greenwald, A. G. (2001). On the malleability of automatic attitudes: Combating automatic prejudice with images of admired and disliked individuals. *Journal of Personality and Social Psychology, 81(5),* 800-814. doi: I0.1037//0022-3514.81.5.800

De Houwer, J. (2002). The implicit association test as a tool for studying dysfunctional associations in psychopathology: strengths and limitations. *Journal of Behavior Therapy and Experimental Psychiatry, 33*, 115-133.

De Young, K. P., Lavender, J. M., Washington, L. A., Looby, A., & Anderson, D. A. (2010). A controlled comparison of the word repeating technique with a word association task. *Journal of Behavior Therapy and Experimental Psychiatry, 41*, 426-432. doi: 10.1016/j.jbtep.2010.04.006

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: An International Quarterly, 25(3),* 218-232. doi: 10.1891/0889-8391.25.3.218

Dobson, K. S., Hollon, S. D., Dimidjian, S., Schmaling, K. B., Kohlenberg, R. J., Gallop,

R. J., Dunner, D. L., & Jacobson, N. S. (2008). Randomized trial of behavioral activation, cognitive therapy, and antidepressant medication in the prevention of relapse and recurrence in major depression. *Journal of Consulting and Clinical Psychology, 76(3),* 468-477. doi: 10.1037/0022-006X.76.3.468

Drake, C. E., Seymour, K. H., & Habib, R. (2016). Testing the IRA: Exploring the reliability and fakability of an idiographic approach to interpersonal attitudes. *The Psychological Record, 66(1)*, 153-163. doi: 10.1007/s40732-015-0160-1

Drake, C. E., Timko, C. A., & Luoma, J. B. (2016). Exploring an implicit measure of acceptance and experiential avoidance of anxiety. *Psychological Record, 66,* 463-475. doi: 10.1007/s40732-016-0186-z

Durham, C. J., McGrath, L. D., Burlingame, G. M., Schaalje, G. B. Lambert, M. J., & Davies, D. R. (2002*). Journal of Psychoeducational Assessment, 20,* 240-257. doi: 10.1177/073428290202000302

Ebert, I. D., Steffens, M. C., Stulpnagel, R. v., & Jelenec, P. (2009). How to like yourself better, or chocolate less: Changing implicit attitudes with one IAT task. *Journal of Experimental Social Psychology, 45,* 1098-1104. doi: 10.1016/j.jesp.2009.06.008

Ellard, K. K., Fairholme, C. P., Boisseau, C. L., Farchione, T. J., & Barlow, D. H. (2010). Unified protocol for the transdiagnostic treatment of emotional disorders: Protocol development and initial outcome data. *Cognitive and Behavioral Practice, 17*, 88-101.

Ellwart, T., Rinck, M., & Becker, E. S. (2006). From fear to love: Individual differences in implicit spider associations. *Emotion, 6(1)*, 18-27. doi: 10.1037/1528-3542.6.1.18

Enrollments by Ethnicity and Race (2015). Southern Illinois University Carbondale

Institutional Research & Studies. Retrieved August 15, 2016, from http://www.irs.siu.edu/quickfacts/students/students_by_race.php

Faul, F., Erdfelder, E., Buchner, A., & Lang, A. G. (2009). Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. Behavior Research Methods, 41(4), 1149-1160. doi:10.3758/BRM.41.4.1149

Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G* Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods, 39(2)*, 175. doi:10.3758/BF03193146

Fleming, P. (2012). Social desirabiloity, not what it seems: A review of the implications for self-reports. *The International Journal of Educational and Psychological Assessment, 11(1),* 3-22.

Gilbert, D. G., McClernon, J., Rabinovich, N. E., Plath, L. C., Masson, C. L., Anderson, A. E., & Sly, K. F. (2002). Mood disturbance fails to resolve across 31 days of cigarette abstinence in women. *Journal of Consulting and Clinical Psychology, 70(1),* 142-152. doi: 10.1037//0022-006X.70.1.142

Gillanders, D. T., Bolderston, H., Bond, F. W., Dempster, M., Flaxman, P. E., Campbell, L., Kerr, S., Tansey, L., Noel, P., Ferenbach, C., Masley, S., Roach, L., Lloyd, J., May, L., Clarke, S., & Remington, B. (2014). The development and initial validation of the cognitive fusion questionnaire. *Behavior Therapy, 45*, 83-101.

Golijani-Moghaddam, N., Hart, A., & Dawson, D. L. (2013). The implicit relational assessment procedure: Emerging reliability and validity data. *Journal of Contextual Behavioral Science, 2,* 105-119. doi: 10.1016/j.jcbs.2013.05.002

Gould, R. A., Buckminster, S., Pollack, M. H., Otto, M. W., Yap, L. (1997). Cognitive-

behavioral and pharmacological treatment for social phobia: A meta-analysis. *Clinical Psychology: Science and Practice, 4(4)*, 291-306.

Greenwald, A. G. & Banaji, M. R. (1995). Implicit social cognition: Attitudes, self-esteem, and stereotypes. *Psychological Review, 102(1),* 4-27.

Greenwald, A. G., McGhee, D. E., & Schwartz, J. L. K. (1998). Measuring individual differences in implicit cognition: The implicit association test. *Journal of Personality and Social Psychology, 74(6)*, 1464-1480.

Greenwald, A. G., Poehlman, T. A., Uhlmann, E. L., & Banaji, M. R. (2009). Understanding and using the implicit association test: III. Meta-analysis of predictive validity. *Journal of Personality and Social Psychology, 97(1)*, 17-41. doi: 10.1037/a0015575

Hanley, C. (1961). Social desirability and response bias in the MMPI. *Journal of Consulting Psychology, 25(1)*, 13-20.

Harris, R. (2009). ACT made simple: An easy-to-read primer on acceptance and commitment therapy. Oakland, CA: New Harbinger Publications, Inc.

Hayes, S. C. (Ed.), Barnes-Holmes, D. (Ed.), & Roche, B. (Ed.) (2001). Relational frame theory: A post-Skinnerian account of human language and cognition. New York, NY: Kluwer Academic/Plenum Publishers.

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*, 1-25. doi:10.1016/j.brat.2005.06.006

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

Hayes, S. C., Strosahl, K. D., & Wilson, K. G. (2012). Acceptance and commitment therapy: The

process and practice of mindful change. New York, NY: Guilford Press.

Healy, H., Barnes-Holmes, Y., Barnes-Holmes, D., & Keogh, C. (2008). An experimental test of a cognitive defusion exercise: Coping with negative and positive self-statements. *The Psychological Record, 58*, 623-640.

Heimberg, R. G., & Barlow, D. H. (1991). New developments in cognitive-behavioral therapy for social phobia. *Journal of Clinical Psychiatry, 52(suppl)*, 21-30.

Hinton, M. J., & Gaynor, S. T. (2010). Cognitive defusion for psychological distress, dysphoria, and low self-esteem: A randomized technique evaluation trial of vocalizing strategies. *International Journal of Behavioral Consultation and Therapy, 6(3)*, 164-185.

Hofmann, S. G., Asnami, A., Vonk, I. J. J., Sayer, A. T., & Fang, A. (2012). The efficacy of cognitive behavioral therapy: A review of meta-analyses. *Cognitive Therapy Research, 36*, 427-440. doi: 10.1007/s10608-012-9476-1

Holtgraves, T. (2004). Social desirability and self-reports: Testing models of socially desirable responding. *Personality and Social Psychology Bulletin, 30(2),* 161-172. doi: 10.1177/0146167203259930

Hooper, N., Villatte, M., Neofotistou, E., & McHugh, L. (2010). The effects of mindfulness versus thought suppression on implicit and explicit measures of experiential avoidance. *International Journal of Behavioral Consultation and Therapy, 6(3),* 233-244.

Hope, D. A., Burns, J. A., Hayes, S. A., Herbert, J. D., & Warner, M. D. (2010). Automatic thoughts and cognitive restructuring in cognitive behavioral group therapy for social anxiety disorder. *Cognitive Therapy and Research, 34*, 1-12. doi: 10.1007/s10608-007-9147-9

Hope, D. A., Heimberg, R. G., & Bruch, M. A. (1995). Dismantling cognitive-behavioral group

therapy for social phobia. *Behavior Research and Therapy, 33(6*), 637-650.

Hussey, I. & Barnes-Holmes, D. (2012). The implicit relational assessment procedure as a measure of implicit depression and the role of psychological flexibility. *Cognitive and Behavioral Practice, 19(4*), 573-582. doi: 10.1016/j.cbpra.2012.03.002

Institute of Medicine (2008). Treatment of posttraumatic stress disorder: An assessment of the evidence. Washington, DC: National Academy Press.

Jacobson, N. S., Dobson, K. S., Truax, P. A., Addis, M. E., Koerner, K., Gollan, J. K., Gortner, E., & Prince, S. E. (2000). A component analysis of cognitive-behavioral treatment for depression. *Journal of Consulting and Clinical Psychology, 64(2),* 295-304.

Karlin, B. E., Walser, R. D., Yesavage, J., Zhang, A., Trockel, M., & Taylor, C. B. (2013). Effectiveness of acceptance and commitment therapy for depression: Comparison among older and younger veterans. *Aging & Mental Health, 17(5),* 555-563. doi: 10.1080/13607863.2013.789002

Kessler, R. C., Petukhova, M., Sampson, N. A., Zaslavsky, A. M., & Wittchen H. U. (2012). Twelve-month and lifetime prevalence and lifetime morbid risk of anxiety and mood disorders in the United States. *International Journal of Methods in Psychiatric Research, 21(3)*, 169-184.

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. doi: 10.1016/j.jcbs.2013.12.001

Kishita, N., Ohtsuki, T., Sakai, M., & Muto, T. (2010). Behavioral assessment methodology for

cognitive defusion: Testing the utility of the implicit relational assessment procedure. *Japanese Journal of Behavior Therapy, 38(3)*, 237-246.

Kuentzel, J. G., Henderson, M. J., & Melville, C. L. (2008). The impact of social desirability biases on self-report among college students and problem gamblers. *Journal of Gambling Studies, 24(3)*, 307-319. doi: 10.1007/s10899-008-9094-8

Leahey, T. M. (2001). A history of modern psychology (3rd ed.). Upper Saddle River, NJ: Prentice Hall, Inc.

Lee, E. B., An, W., Levin, M. E., & Twohig, M. P. (2015). An initial meta-analysis of acceptance and commitment therapy for treating substance use disorders. *Drug and Alcohol Dependence, 155*, 1-7. http://dx.doi.org/10.1016/j.drugalcdep.2015.08.004

Lindgren, K. P., Neighbors, C., Teachman, B. A., Wiers, R. W., Westgate, E., & Greenwald, A. G. (2013). I drink therefore I am: Validating alcohol-related implicit association tests. *Psychology of Addictive Behaviors, 27(1)*, 1-13. doi: 10.1037/a0027640

Longmore, R. J., & Worrell, M. (2007). Do we need to challenge thoughts in cognitive behavior therapy? *Clinical Psychology Review, 27(2),* 173–187. doi: 10.1016/j.cpr.2006.08.001

Lovibond, P. F., & Lovibond, S. H. (1995). The structure of negative emotional states: Comparison of the depression anxiety stress scales (DASS) with the Beck depression and anxiety inventories. *Behaviour Research and Therapy, 33(3)*, 335-343. http://dx.doi.org.proxy.lib.siu.edu/10.1016/0005-7967(94)00075-U

Maclean, J. C., Xu, H., French, M. T., & Ettner, S. L. (2014). Mental health and high-cost health care utilization: New evidence from axis II disorders. *Health Services Research, 49(2),* 683-704. doi: 10.1111/1475-6773.12107

Masuda, A., Feinstein, A. B., Wendell, J. W., & Sheehan, S. T. (2010). Cognitive defusion

versus thought distraction: A clinical rationale, training, and experiential exercise in altering psychological impacts of negative self-referential thoughts. *Behavior Modification, 34(6),* 520-538. doi: 10.1177/0145445510379632

Masuda, A., Hayes, S. C., Sackett, C. F., & Twohig, M. P. (2004). Cognitive defusion and self-relevant negative thoughts: examining the impact of a ninety year old technique. *Behaviour Research and Therapy, 42*, 477-485. doi: 10.1016/j.brat.2003.10.008

Masuda, A., Hayes, S. C., Twohig, M. P., Drossel, C., Lillis, J., & Washio, Y. (2009). A parametric study of cognitive defusion and the believability and discomfort of negative self-relevant thoughts. *Behavior Modification, 33(2),* 250-262. doi: 10.1177/0145445508326259

Masuda, A., Twohig, M. P., Stormo, A. R., Feinstein, A. B., Chou, 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*, 11-17. doi: 10.1016/j.jbtep.2009.08.006

Mattick, R. P., & Peters, L. (1988). Treatment of severe social phobia: Effects of guided exposure with and without cognitive restructuring. *Journal of Consulting and Clinical Psychology, 56(2),* 251-260.

McConnell, A. R. & Leibold, J. M. (2001). Relations among the implicit association test, discriminatory behavior, and explicit measures of racial attitudes. *Journal of Experimental Social Psychology, 37(5),* 435-442. doi: 10.1006/jesp.2000.1470

Moffitt, R., Brinkworth, G., Noakes, M., & Mohr, P. (2012). A comparison of cognitive restructuring and cognitive defusion as strategies for resisting craved food. *Psychology & Health, 27,* 74-90. http://dx.doi.org/10.1080/08870446.2012.694436

Nicholson, E. & Barnes-Holmes, D. (2012). The implicit relational assessment procedure (IRAP) as a measure of spider fear. *The Psychological Record, 62*, 263-278.

Nosek, B.A., Greenwald, A. G., & Banaji, M. R. (2007).The implicit association test at age 7: A methodological and conceptual review. In: J. A. Bargh (Ed.), *Automatic processes in social thinking and behavior* (pp. 265–292). London: Psychology Press.

Osman, A., Wong, J. L., Bagge, C. L., Freedenthal, S., Gutierrez, P. M., & Lozano, G. (2012). The depression anxiety stress scales- 21 (DASS-21): further examination of dimesions, scale reliability, and correlates. *Journal of Clinical Psychology, 68(12)*, 1322-1338. doi: 10.1002/jclp.21908

Ritzert, T. R., Forsyth, J. P., Berghoff, C. R., Barnes-Holmes, D., & Nicholson, E. (2015). The impact of a cognitive defusion intervention on behavioral and psychological flexibility: An experimental evaluation in a spider fearful non-clinical sample. *Journal of Contextual Behavioral Science, 4,* 112-120. http://dx.doi.org/10.1016/j.jcbs.2015.04.001

Roddy, S., Stewart, I., & Barnes-Holmes, D. (2011). Facial reactions reveal that slim is good but fat is not bad: Implicit and explicit measures of body-size bias. *European Journal of Social Psychology, 41,* 688-694.

Rodriguez, B. F., Weisberg, R. B., Pagano, M. E., Machan, J. T., Culpepper, L., & Keller, M. B. (2004). Frequency and patterns of psychiatric comorbidity in a sample of primary care patients with anxiety disorders. *Comprehensive Psychiatry, 45(2),* 129-137. doi: 10.1016/j.comppsych.2003.09.005

Rolffs, J. L., Rogge, R. D., & Wilson, K. G. (2016). Disentangling components of flexibility via the hexaflex model: Development and validation of the multidimensional psychological flexibility inventory (MPFI). *Assessment,* 1-25. doi: 10.1177/1073191116645905

Salaberria, K., & Echeburua, E. (1998). Long-term outcome of cognitive therapy's contribution to self-exposure in vivo to the treatment of generalized social phobia. *Behavior Modification, 22(3*), 262-284. doi: 10.1177/01454455980223003

Schofield, D. J., Shrestha, R. N., Percival, R., Passey, M. E., Callander, E. J., & Kelly, S. J. (2011). The personal and national costs of mental health conditions: impacts of income, taxes, government support payments due to lost labour force participation. *BMC Psychiatry, 11:72*. http://www.biomedcentral.com/1471-244X/11/72

Scholing, A., & Emmelkamp, P. M. G. (1996). Treatment of generalized social phobia: Results at long-term follow-up. *Behavior Research and Therapy, 34(5),* 447-452.

Sharpe, J. P. & Gilbert, D. G. (1998). Effects of repeated administration of the Beck depression inventory and other measures of negative mood states. *Personality and Individual Differences, 24(4),* 457-463.

Taylor, S., Woody, S., Koch, W. J., McLean, P., Paterson, R. J., & Anderson, K. W. (1997). Cognitive restructuring in the treatment of social phobia. *Behavior Modification, 21(4),* 487-511. doi: 10.1177/01454455970214006

Tabachnick, B. G. & Fidell, L. S. (2013). Using multivariate statistics (6th ed.). Upper Saddle River, NJ: Pearsons Education, Inc.

Teachman, B. A. & Woody, S. R. (2003). Automatic processing in spider phobia: Implicit fear associations over the course of treatment. *Journal of Abnormal Psychology, 112(1),* 100-109. doi: 10.1037/0021-843X.112.1.100

Titchener, E. B. (1916). A textbook of psychology. New York: MacMillan.

Twohig, M. P., Hayes, S. C., Plumb, J. C., Pruitt, L. D., Collins, A. B., Hazlett-Stevens, H., &

Woidneck, M. R. (2010). A randomized clinical trial of acceptance and commitment therapy versus progressive relaxation training for obsessive-compulsive disorder. *Journal of Consulting and Clinical Psychology, 78(5)*, 705-716. doi: 10.1037/a0020508

Vahey, N. A., Barnes-Holmes, D. Barnes-Holmes, Y., & Stewart, I. (2009). A first test of the implicit relational assessment procedure as a measure of self-esteem: Irish prisoner groups and university students. *The Psychological Record, 59,* 371-388.

Vahey, N. A., Nicholson, E., & Barnes-Holmes, D. (2015). A meta-analysis of criterion effects for the implicit relational assessment procedure (IRAP) in the clinical domain. *Journal of Behavior Therapy and Experimental Psychiatry, 48,* 59-65. http://dx.doi.org.proxy.lib.siu.edu/10.1016/j.jbtep.2015.01.004

van Dam-Baggen, R., & Kraaimaat, F. (2000). Group social skills training or cognitive group therapy as the clinical treatment of choice for generalized social phobia? *Journal of Anxiety Disorders, 14(5),* 437-451.

Vecchione, M., Dentale, F., Alessandri, G., & Barbaranelli, C. (2014). Fakability of implicit and explicit measures of the big five: Research findings from organizational settings. *International Journal of Selection and Assessment, 22(2)*, 211-218.

Walser, R. D., Karlin, B. E., Trockel, M., MAzine, B. & Taylor, C. B. (2013). Training in and implementation of acceptance and commitment therapy for depression in the veterans health administration: Therapist and patient outcomes. *Behaviour Research and Therapy, 51*, 555-563. doi: 10.1016/j.brat.2013.05.009

Wolitzky-Taylor, K. B., Arch, J. J., Rosenfield, D., & Craske, M. G. (2012). Moderators and

non-specific predictors of treatment outcome for anxiety disorders: A comparison of cognitive behavioral therapy to acceptance and commitment therapy. *Journal of Consulting and Clinical Psychology, 80(5),* 786-799. doi: 10.1037/a0029418

Zhao, W., Zhou, Y., Liu, X-P., & Ran, L-W. (2013). Effectiveness of acceptance and commitment therapy on depression. *Chinese Journal of Clinical Psychology, 21(1)*, 153-157.

Ziegler, M. (2015).” F*** you, I won’t do what you told me!” – Response biases as threats to psychological assessment. *European Journal of Psychological Assessment, 31(3)*, 153-158. doi: 10.1027/1015-5759/a000292

APPENDICES

Appendix A

Demographics Questionnaire

Age (in years): _____
Are you a member of a sorority or fraternity?
- Yes
- No

Country of Origin (the country you regard as your home):
- United States
- Other

English as First Language (select whether or not English is your first language):
- Yes
- No

Education (select your current status as a student):
- freshman
- sophomore
- junior
- senior

Declared Major:
- Psychology
- Undecided/Undeclared
- Other (please specify): _____

Political Affiliation (select the party that you most identify with):
- Democrat
- Republican
- Other (please specify): _____

Race/Ethnicity (select as many as are appropriate for you):
- American Indian or Alaska Native
- Asian
- Black or African-American
- Hispanic or Latino
- Native Hawaiian or Other Pacific Islander
- White or Caucasian
- Other

Religion (select the category that you most identify with):
- Agnostic (undecided as to the existence of God or an afterlife)
- Atheist (do not believe in the existence of God or an afterlife)
- Buddhist
- Christian (any denomination of Catholics, Protestants, etc.)
- Hindu
- Jewish
- Muslim
- Other (please specify): _____

Gender:
- Female

Male
Other

Sexual Identity:

Bisexual (attracted to both sexes)
Heterosexual (attracted to the opposite sex)
Homosexual (attracted to the same sex)

Socioeconomic Status (if someone other than you is providing more than 50% of your income, please report his or her annual income instead):

$25,000 or less
$25,001-$50,000
$50,001-$75,000
$75,001 or more

Therapy Experience (select whether or not you have ever participated in psychotherapy):

Yes
No

Food Allergies (indicate if you have any food allergies or illness such as lactose intolerance, Celiac Disease, nut allergy, etc.):

Yes (please specify)_____
No

Appendix B

Cognitive Fusion Questionnaire

CFQ

Below you will find a list of statements. Please rate how true each statement is for you by circling a number next to it. Use the scale below to make your choice.

1	2	3	4	5	6	7
never true	very seldom true	seldom true	sometimes true	frequently true	almost always true	always true

1.	My thoughts cause me distress or emotional pain	1 2 3 4 5 6 7
2.	I get so caught up in my thoughts that I am unable to do the things that I most want to do	1 2 3 4 5 6 7
3.	I over-analyse situations to the point where it's unhelpful to me	1 2 3 4 5 6 7
4.	I struggle with my thoughts	1 2 3 4 5 6 7
5.	I get upset with myself for having certain thoughts	1 2 3 4 5 6 7
6.	I tend to get very entangled in my thoughts	1 2 3 4 5 6 7
7.	It's such a struggle to let go of upsetting thoughts even when I know that letting go would be helpful	1 2 3 4 5 6 7

Thank you for completing this questionnaire

Appendix C

Depression Anxiety and Stress Scale 21

Please read each statement and circle a number 0, 1, 2 or 3 which indicates how much the statement applied to you over the past week. There are no right or wrong answers. Do not spend too much time on any statement.

The rating scale is as follows:

0 Did not apply to me at all - NEVER

1 Applied to me to some degree, or some of the time - SOMETIMES

2 Applied to me to a considerable degree, or a good part of time - OFTEN

3 Applied to me very much, or most of the time - ALMOST ALWAYS

		N	S	O	AA	FOR OFFICE USE D	A	S
1	I found it hard to wind down	0	1	2	3			
2	I was aware of dryness of my mouth	0	1	2	3			
3	I couldn't seem to experience any positive feeling at all	0	1	2	3			
4	I experienced breathing difficulty (eg, excessively rapid breathing, breathlessness in the absence of physical exertion)	0	1	2	3			
5	I found it difficult to work up the initiative to do things	0	1	2	3			
6	I tended to over-react to situations	0	1	2	3			
7	I experienced trembling (eg, in the hands)	0	1	2	3			
8	I felt that I was using a lot of nervous energy	0	1	2	3			
9	I was worried about situations in which I might panic and make a fool of myself	0	1	2	3			
10	I felt that I had nothing to look forward to	0	1	2	3			
11	I found myself getting agitated	0	1	2	3			
12	I found it difficult to relax	0	1	2	3			
13	I felt down-hearted and blue	0	1	2	3			
14	I was intolerant of anything that kept me from getting on with what I was doing	0	1	2	3			
15	I felt I was close to panic	0	1	2	3			
16	I was unable to become enthusiastic about anything	0	1	2	3			
17	I felt I wasn't worth much as a person	0	1	2	3			
18	I felt that I was rather touchy	0	1	2	3			
19	I was aware of the action of my heart in the absence of physicalexertion (eg, sense of heart rate increase, heart missing a beat)	0	1	2	3			
20	I felt scared without any good reason	0	1	2	3			
21	I felt that life was meaningless	0	1	2	3			
					TOTALS			

Appendix D

Food Attitudes Questionnaire

Please indicate how many days of the week you drink the following:								
	0	1	2	3	4	5	6	7
milk								
water								
pop/soda								
coffee								

Please indicate how much you **like** drinking the following:							
	-3 not at all	-2	-1	0 neutral	1	2	3 very much
milk							
Water							
Pop/soda							
coffee							

Please indicate how many days of the week you eat the following:								
	0	1	2	3	4	5	6	7
apples								
lemons								
oranges								
bananas								

Please indicate how much you **like** eating the following:							
	-3 not at all	-2	-1	0 neutral	1	2	3 very much
apples							
lemons							
oranges							
bananas							

Appendix E

Milk-Lemon Defusion Measure

Please indicate how relevant each of the following words are when you read the word "milk":											
	0 not at all	1	2	3	4	5	6	7	8	9	10 extremely
creamy											
white											
cow											
dairy											
word											
text											
sound											
symbol											

Please indicate how relevant each of the following words are when you read the word "lemon":											
	0 not at all	1	2	3	4	5	6	7	8	9	10 extremely
yellow											
fruit											
citrus											
juicy											
word											
text											
sound											
symbol											

Appendix F

Multidimensional Psychological Flexibility Inventory

FLEXIBILITY SUBSCALES

ACCEPTANCE

IN THE LAST TWO WEEKS…	Never TRUE	Rarely TRUE	Occasionally TRUE	Often TRUE	Very Often TRUE	Always TRUE
I was receptive to observing unpleasant thoughts and feelings without interfering with them.	O	O	O	O	O	O
I tried to make peace with my negative thoughts and feelings rather than resisting them	O	O	O	O	O	O
I made room to fully experience negative thoughts and emotions, breathing them in rather than pushing them away	O	O	O	O	O	O
When I had an upsetting thought or emotion, I tried to give it space rather than ignoring it	O	O	O	O	O	O
I opened myself to all of my feelings, the good and the bad	O	O	O	O	O	O

PRESENT MOMENT AWARENESS

IN THE LAST TWO WEEKS…	Never TRUE	Rarely TRUE	Occasionally TRUE	Often TRUE	Very Often TRUE	Always TRUE
I was attentive and aware of my emotions	O	O	O	O	O	O
I was in tune with my thoughts and feelings from moment to moment	O	O	O	O	O	O
I paid close attention to what I was thinking and feeling	O	O	O	O	O	O

I was in touch with the ebb and flow of my thoughts and feelings	O	O	O	O	O	O
I strived to remain mindful and aware of my own thoughts and emotions	O	O	O	O	O	O

SELF AS CONTEXT

IN THE LAST TWO WEEKS…	**Never TRUE**	**Rarely TRUE**	**Occasionally TRUE**	**Often TRUE**	**Very Often TRUE**	**Always TRUE**
Even when I felt hurt or upset, I tried to maintain a broader perspective	O	O	O	O	O	O
I carried myself through tough moments by seeing my life from a larger viewpoint	O	O	O	O	O	O
I tried to keep perspective even when life knocked me down	O	O	O	O	O	O
When I was scared or afraid, I still tried to see the larger picture	O	O	O	O	O	O
When something painful happened, I tried to take a balanced view of the situation	O	O	O	O	O	O

DEFUSION

IN THE LAST TWO WEEKS…	**Never TRUE**	**Rarely TRUE**	**Occasionally TRUE**	**Often TRUE**	**Very Often TRUE**	**Always TRUE**
I was able to let negative feelings come and go without getting caught up in them	O	O	O	O	O	O
When I was upset, I was able to let those negative feelings pass through me without clinging to them	O	O	O	O	O	O
When I was scared or afraid, I was able to gently experience those feelings, allowing them to pass	O	O	O	O	O	O
I was able to step back and notice negative thoughts and feelings without reacting to them	O	O	O	O	O	O
In tough situations, I was able to notice my thoughts and feelings without getting overwhelmed by them	O	O	O	O	O	O

VALUES

IN THE LAST TWO WEEKS…	Never TRUE	Rarely TRUE	Occasionally TRUE	Often TRUE	Very Often TRUE	Always TRUE
I was very in-touch with what is important to me and my life	O	O	O	O	O	O
I stuck to my deeper priorities in life	O	O	O	O	O	O
I tried to connect with what is truly important to me on a daily basis	O	O	O	O	O	O
Even when it meant making tough choices, I still tried to prioritize the things that were important to me	O	O	O	O	O	O
My deeper values consistently gave direction to my life	O	O	O	O	O	O

COMMITTED ACTION

IN THE LAST TWO WEEKS…	Never TRUE	Rarely TRUE	Occasionally TRUE	Often TRUE	Very Often TRUE	Always TRUE
Even when I stumbled in my efforts, I didn't quit working toward what is important	O	O	O	O	O	O
Even when times got tough, I was still able to take steps toward what I value in life	O	O	O	O	O	O
Even when life got stressful and hectic, I still worked toward things that were important to me	O	O	O	O	O	O
I didn't let set-backs slow me down in taking action toward what I really want in life	O	O	O	O	O	O
I didn't let my own fears and doubts get in the way of taking action toward my goals	O	O	O	O	O	O

INFLEXIBILITY SUBSCALES

EXPERIENTIAL AVOIDANCE

IN THE LAST TWO WEEKS…	Never TRUE	Rarely TRUE	Occasionally TRUE	Often TRUE	Very Often TRUE	Always TRUE

When I had a bad memory, I tried to distract myself to make it go away	O	O	O	O	O	O
I tried to distract myself when I felt unpleasant emotions	O	O	O	O	O	O
When unpleasant memories came to me, I tried to put them out of my mind	O	O	O	O	O	O
When something upsetting came up, I tried very hard to stop thinking about it	O	O	O	O	O	O
If there was something I didn't want to think about, I would try many things to get it out of my mind	O	O	O	O	O	O

LACK OF CONTACT WITH THE PRESENT MOMENT

IN THE LAST TWO WEEKS…	**Never TRUE**	**Rarely TRUE**	**Occasionally TRUE**	**Often TRUE**	**Very Often TRUE**	**Always TRUE**
I did most things on "automatic" with little awareness of what I was doing.	O	O	O	O	O	O
I did most things mindlessly without paying much attention.	O	O	O	O	O	O
I went through most days on auto-pilot without paying much attention to what I was thinking or feeling	O	O	O	O	O	O
I floated through most days without paying much attention.	O	O	O	O	O	O
Most of the time I was just going through the motions without paying much attention	O	O	O	O	O	O

SELF AS CONTENT

IN THE LAST TWO WEEKS…	**Never TRUE**	**Rarely TRUE**	**Occasionally TRUE**	**Often TRUE**	**Very Often TRUE**	**Always TRUE**
I thought some of my emotions were bad or inappropriate and I shouldn't feel them	O	O	O	O	O	O
I criticized myself for having irrational or inappropriate emotions	O	O	O	O	O	O

I believed some of my thoughts are abnormal or bad and I shouldn't think that way	O	O	O	O	O	O
I told myself that I shouldn't be feeling the way I'm feeling	O	O	O	O	O	O
I told myself I shouldn't be thinking the way I was thinking	O	O	O	O	O	O

FUSION

IN THE LAST TWO WEEKS…	**Never TRUE**	**Rarely TRUE**	**Occasionally TRUE**	**Often TRUE**	**Very Often TRUE**	**Always TRUE**
Negative thoughts and feelings tended to stick with me for a long time.	O	O	O	O	O	O
Distressing thoughts tended to spin around in my mind like a broken record.	O	O	O	O	O	O
It was very easy to get trapped into unwanted thoughts and feelings.	O	O	O	O	O	O
When I had negative thoughts or feelings it was very hard to see past them.	O	O	O	O	O	O
When something bad happened it was hard for me to stop thinking about it.	O	O	O	O	O	O

LACK OF CONTACT WITH VALUES

IN THE LAST TWO WEEKS…	**Never TRUE**	**Rarely TRUE**	**Occasionally TRUE**	**Often TRUE**	**Very Often TRUE**	**Always TRUE**
My priorities and values often fell by the wayside in my day to day life	O	O	O	O	O	O
When life got hectic, I often lost touch with the things I value	O	O	O	O	O	O
The things that I value the most often fell off my priority list completely	O	O	O	O	O	O

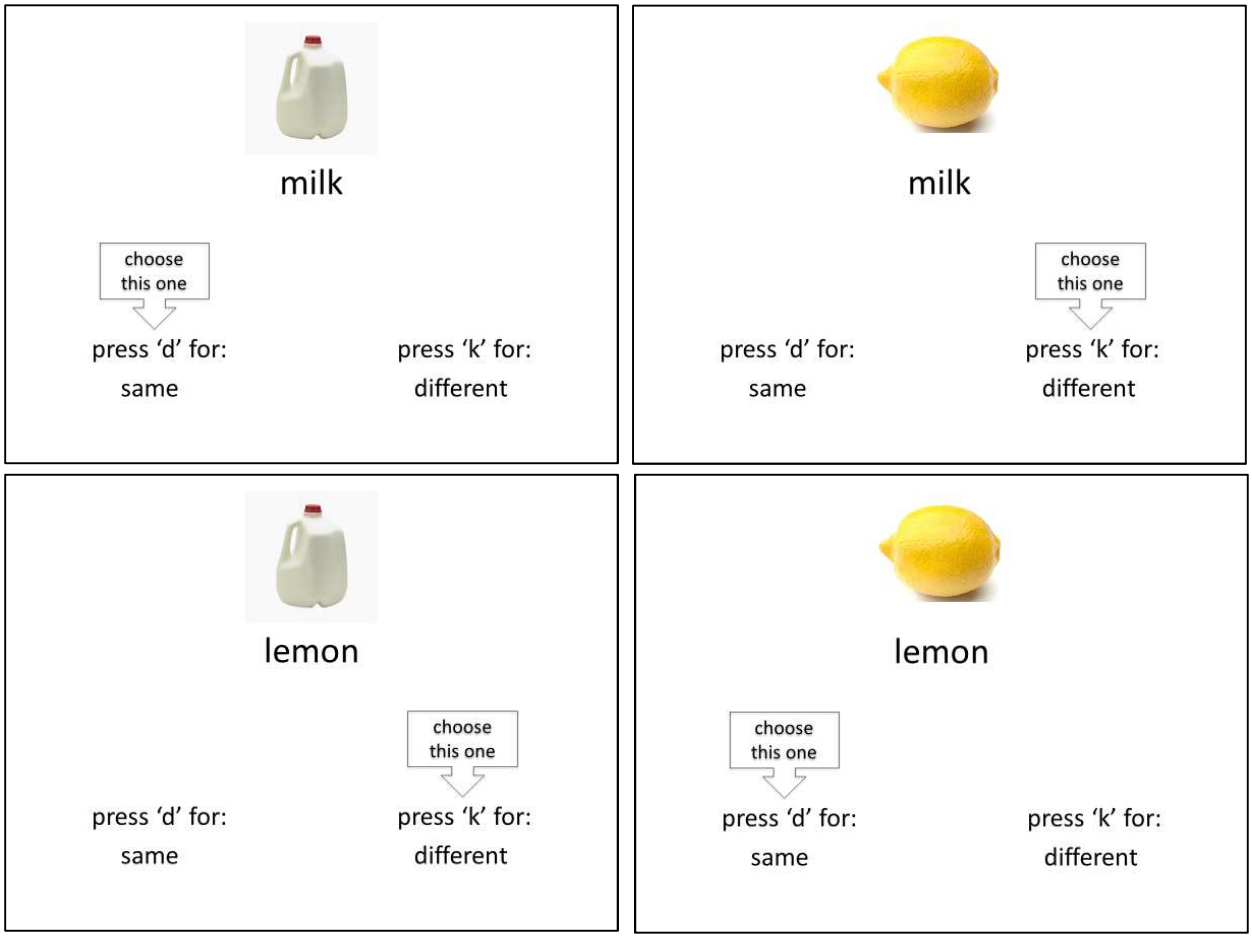
I didn't usually have time to focus on the things that are really important to me	O	O	O	O	O	O
When times got tough, it was easy to forget about what I truly value	O	O	O	O	O	O

INACTION

IN THE LAST TWO WEEKS…	**Never TRUE**	**Rarely TRUE**	**Occasionally TRUE**	**Often TRUE**	**Very Often TRUE**	**Always TRUE**
Negative feelings often trapped me in inaction	O	O	O	O	O	O
Negative feelings easily stalled out my plans	O	O	O	O	O	O
Getting upset left me stuck and inactive	O	O	O	O	O	O
Negative experiences derailed me from what's really important	O	O	O	O	O	O
Unpleasant thoughts and feelings easily overwhelmed my efforts to deepen my life	O	O	O	O	O	O

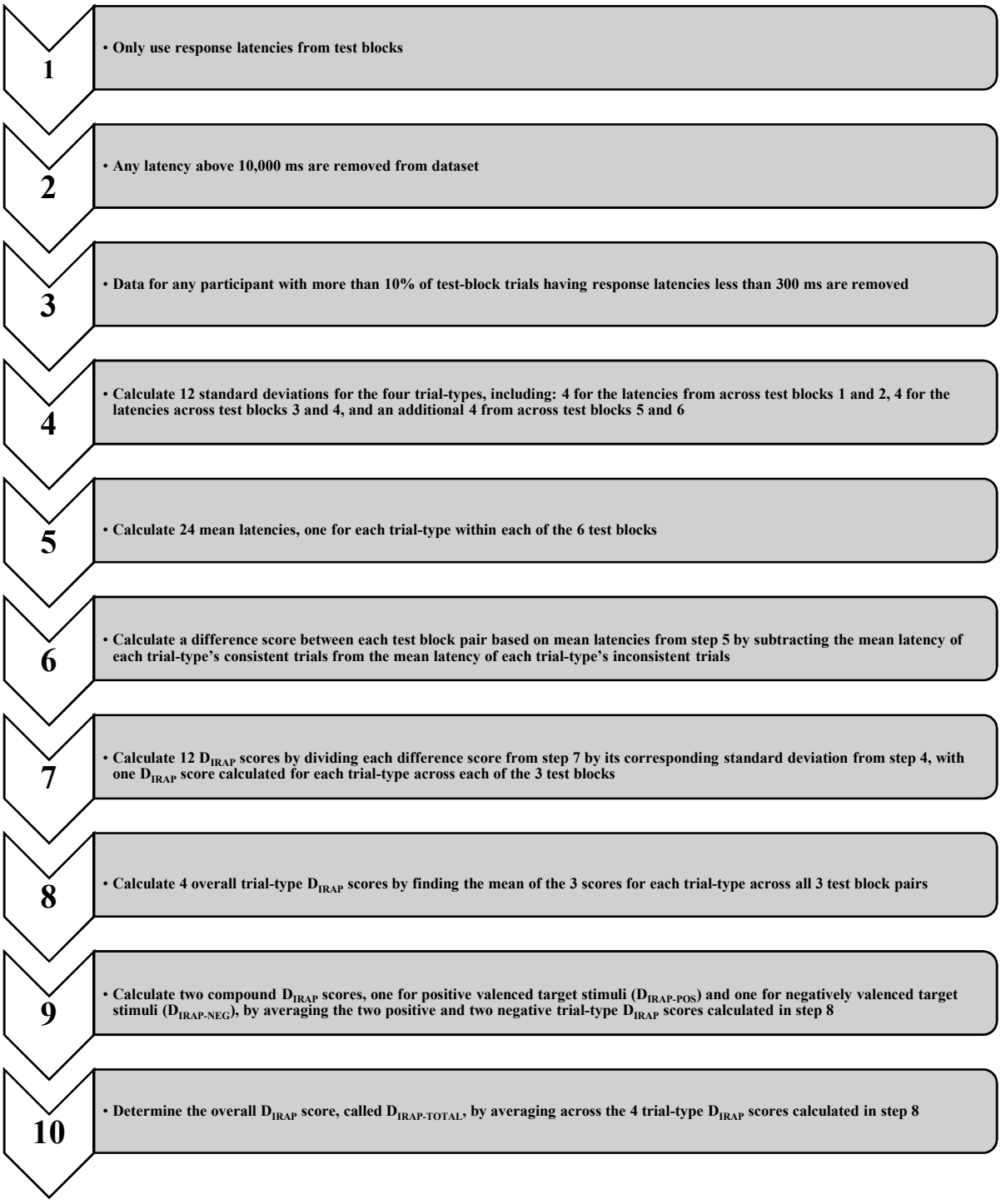
Appendix G

IRAP Trial-Types



Appendix H

Procedure for Calculating *D*-IRAP Scores



Appendix I

Informed Consent

The objective of this study is to examine the effects of a psychological intervention on a computerized task that may be a useful measure of behavior. More specifically, we want to investigate if completing a psychological intervention will result in differential effects on the Implicit Relational Assessment Procedure (IRAP).

I understand that as a participant in this study, I will be asked to complete a computer task and a demographics questionnaire. I understand that the study investigator is mandated to report any intention on my part to harm myself. It is possible that I may find parts of the task uncomfortable and I may refuse to answer or withdraw from the study at any time without penalty. As a participant in this study, I agree to complete the computer task and the questionnaires. If I have any questions about this study, I may contact Travis Sain at travis.sain@siu.edu for more information.

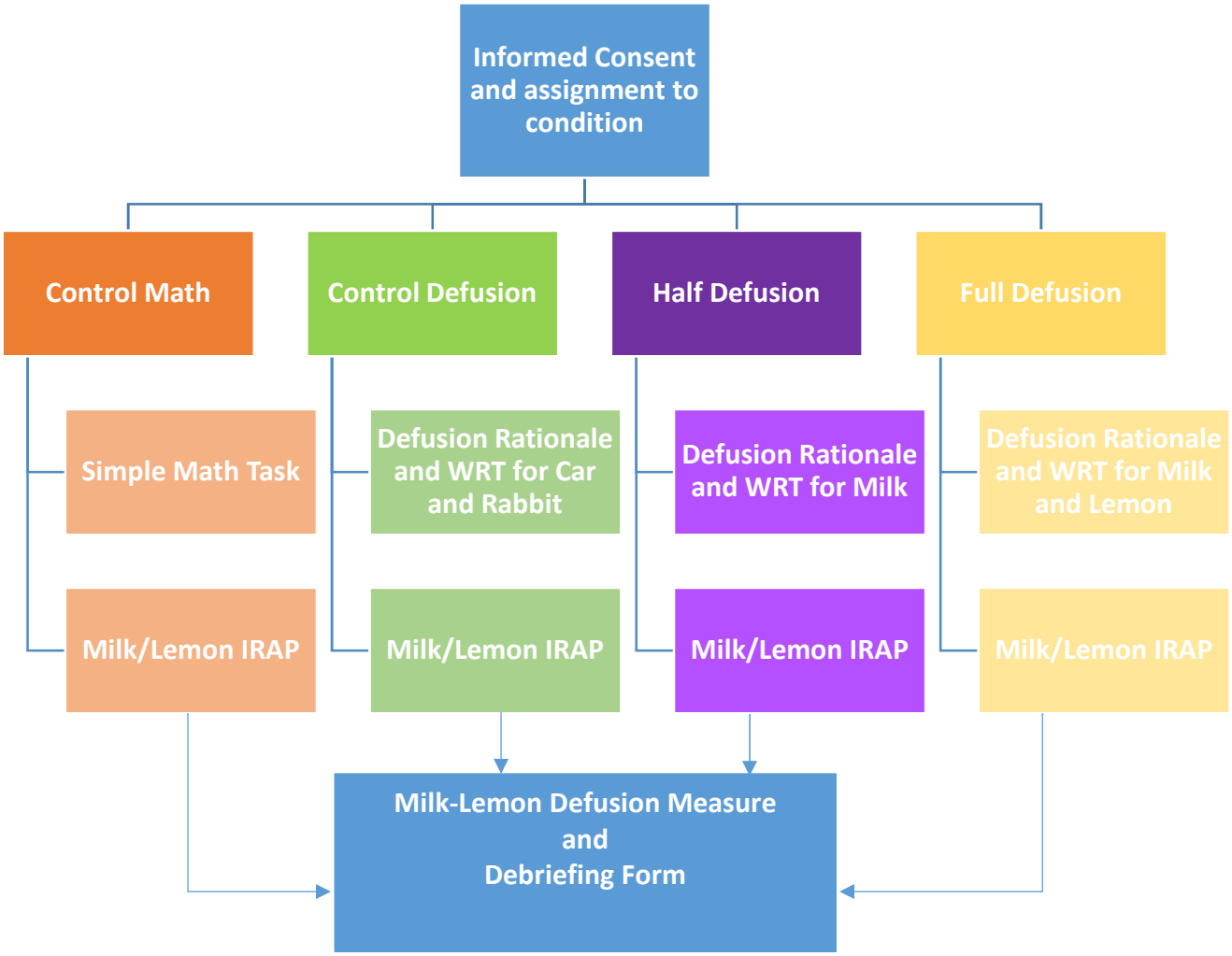
I understand that my participation in this research is voluntary and that I may withdraw from the study at any time, without penalty. This study will require approximately 60 minutes of my time. For my participation, I will receive 4 credits. Furthermore, I understand that all material received from my participation will be kept confidential and that my name/identity will in no way be connected with my answers. Instead, only an assigned subject number will be used in association with my answers.

I have read and understand the information above,

Signature Date

Appendix J

Flowchart of Procedures



Appendix K

Debriefing Form

You have just completed a study involving the Implicit Relational Assessment Procedure (IRAP). The study investigators are interested in examining the effects of a psychological intervention known as defusion on the IRAP so that we may be gain a better understanding of how defusion and the IRAP work. In order to establish the usefulness of this measure, we need to administer the IRAP along with other measures so that we can understand how people react to the measure.

We appreciate your willingness to contribute to our efforts to understand defusion and the IRAP. If you have any additional questions about this study, please contact Travis Sain at travis.sain@siu.edu.

This project has been reviewed and approved by the SIUC Human Subjects Committee. Questions concerning your rights as a participant in this research may be addressed to the Committee Chairperson, Office of Sponsored Projects Administration, SIUC, Carbondale, IL 62901-4709. Phone (618) 453-4533. E-mail: siuhsc@siu.edu

Appendix L

Defusion Exercise Script

Note: The Full Defusion exercise is conducted by reading all text, while the Half Defusion exercise is conducted by omitting the **bold** text.

During the next few minutes I want to teach you a psychological skill known as defusion. Defusion is a skill that we apply to forms of language, such as written or spoken words, thoughts and beliefs, and rules that people follow in their lives.

I am interested in defusion because much of the time people are fused with their thoughts and beliefs, and this can lead to psychological problems. Fusion means we take words literally – we believe our thoughts, obey our thoughts, and try to fix our thoughts if we don't like them. Fusion means that the words that go through our mind are taken seriously. When we are fused, the meaning of words may generate emotional reactions and also may control our behavior. This is not always a good thing. For example, if you have the thought, "I'm worthless", this might affect your mood and might affect your actual behavior, especially if you believe the thought, which is a sign of fusion.

Here's a really simple example of fusion. I am going to say a word, and as quickly as you can, I want you to tell me what goes through your mind after I say it. Is that okay? Are you willing? Okay, here's the word: milk.

[Spend 10-20 seconds noting the participant's reactions to the word.]

Okay, let's try another one. I am going to say another word, and I want you to tell me what goes through your mind after I say it: lemon.

[Spend 10-20 seconds noting the participant's reactions to the word.]

Isn't this interesting? There is no actual milk in this room. **There's also no lemons in this room.** All I did was make a **couple of** sound**s** with my mouth – "milk" **and "lemon"**, and all these things popped into your mind. The reason your mind generated all of that stuff is because you have learned to add a lot of meaning and symbolism to words, to sounds that you hear from others and to lines of text that you can see on a piece of paper or computer screen.

Okay, now I want to show you a quick way to defuse a word. I want you to repeat the word "milk" with me as quickly as you can for the next 30 seconds. As you are repeating the word over and over all I want you to do is notice if anything about the word changes for you as you complete the exercise. It might be a little weird to do this, but that's okay because it's supposed to be a little weird. Is that okay? Are you willing to repeat the word with me? Okay, let's go.

[Start a timer and begin repeating "milk" with the participant. After 10 seconds ask them to say the word faster. After 20 seconds ask them to say the word louder. After 30 seconds has passed put your hand up and say "Okay, you can stop."]

Thanks for doing that. What was that like? Did you notice anything about your perception of the word as we repeated it?

[Spend about 30 seconds discussing the participant's observations. Listen for any comments indicating that the meaning of the word, and/or the participant's initial reactions to the word,

reduced. Also listen for any comments indicating that awareness of basic auditory and/or motor sensations that were present while repeating the word increased.]

Okay, let's try the same thing with the word "lemon". Let's repeat it as quickly as we can for 30 seconds. Are you ready?

[Start a timer and begin repeating "lemon" with the participant. After 10 seconds ask them to say the word faster. After 20 seconds ask them to say the word louder. After 30 seconds has passed put your hand up and say "Okay, you can stop."]

So, how did that go? What did you notice?

[Spend about 20 seconds discussing the participant's observations. Listen for any comments indicating that the meaning of the word, and/or the participant's initial reactions to the word, reduced. Also listen for any comments indicating that awareness of basic auditory and/or motor sensations that were present while repeating the word increased.]

Okay, so this was a small example of what it is like to experience defusion, in this case with the word**s** "milk" **and "lemon"**. In a very fundamental way, all words are just sounds in the air and just squiggly lines on paper – and on a computer screen. So, during the next part of the study, I want you to try to stay aware of this way of seeing **these**/this word**s**.

VITA

Graduate School
Southern Illinois University

Travis Sain

travis.n.sain@gmail.com

Grand Valley State University
Bachelor of Science, Psychology, May 2011

Southern Illinois University Carbondale
Master of Arts, Clinical Psychology, August 2015

Dissertation Title:

Tasting what you see: Using the implicit relational assessment procedure to measure the effects of the word repetition technique

Major Professor: Chad E. Drake

Publications:

Lakey, B., & Sain, T. (in press). The ordinary conversation scale. To appear in D. L. Worthington and G. Bodie (Eds), Sourcebook of Listening Methodology and Measurement. Wiley.

Woods, W. C., Lakey, B., & Sain, T. (2016). The role of ordinary conversation and shared activity in the main effect between perceived support and affect. *European Journal of Social Psychology, 46*, 356-368.

Drake, C. E., Kramer, S., Sain, T., Swiatek, R., Kohn, K., & Murphy, M. (2015). Exploring the reliability and convergent validity of implicit racial evaluations. *Behavior and Social Issues, 24,* 68-87.