

Cognitive and Emotional Defusion:  
Investigations of the Underlying Psychological Processes  
using Explicit and Implicit Measures



**NUI MAYNOOTH**  
Coláiste na hÉireann Mhá Nuall



**IRCHSS**

Claire Keogh, B.A. (Psychology)

Thesis submitted to the National University of Ireland,  
Maynooth in fulfilment of the requirements for a Ph.D. in  
Psychology

Supervised by Dr. Yvonne Barnes-Holmes  
Department of Psychology  
October 2008

## CONTENTS

	<b>Page</b>	
Title	<i>i</i>	
Contents	<i>ii</i>	
Acknowledgements (Academic)	<i>iv</i>	
Acknowledgements (Personal)	<i>v</i>	
Abstract	<i>vi</i>	
List of Tables	<i>viii</i>	
List of Figures	<i>x</i>	
Chapter 1	General Introduction	1
<b>PART I</b>		
Chapter 2	An Experimental Test of a Brief Cognitive Defusion Exercise: Coping with Negative and Positive Self-Statements	29
Chapter 3	Cognitive Defusion and Personally Relevant Negative Thoughts: Exploring the Relative Impact of Therapeutic Instructions and Experiential Exercise	61
<b>PART II</b>		
Chapter 4	Developing an Implicit Measure of Reactions to Positive and Negative Self-Relevant Words using the Implicit Relational Assessment Procedure (IRAP)	92
Chapter 5	Using the IRAP to Measure Reactions to Positive and Negative Self- Statements: Comparing Undergraduate Students with Acceptance and Commitment Therapy (ACT) Therapists	142
Chapter 6	Using the IRAP to Measure Reactions To Positive and Negative Self-Statements Following a Cognitive Defusion Intervention	166

Chapter 7	General Discussion	183
References		208
Appendices		229

## **Academic Acknowledgements**

This research was funded by the Irish Research Council for the Humanities and the Social Sciences, through a Government of Ireland Scholarship.

Special thanks to Dermot for allowing us to use the Implicit Relational Assessment Procedure. Thanks also for all your advice on using this measure and for providing invaluable advice at all times throughout the years!

Thanks to Akihiko Masuda, who kindly sent us copies of the manuscripts he used in previous research (Masuda, Hayes, Sackett, & Twohig, 2004).

Thanks also to a number of anonymous reviewers who provided constructive comments and advice on the thesis work submitted for publication to date.

Specifically, the work presented in Chapter 1 has now been published in *The Irish Psychologist*, while the experimental work in Chapter 2 has been accepted for publication in *The Psychological Record*.

Finally, thanks to Derek, Anne and Victoria who helped out in every way imaginable over the last seven years.

## **Personal Acknowledgements**

Yvonne - Thank you for your time, patience, knowledge, support and kindness. You are most generous in every way imaginable. Thank you for everything over the last few years. Wow! I will never forget this experience!

To Mam and Dad - Thanks for always encouraging me. I can't imagine surviving the last few months without you both there to love and support me when I needed it most.

To all my friends from NUIM, psychs and honorary-psychs alike! Lunchtimes will *really* never be the same again!

## Abstract

The aim of the current thesis was to conduct experimental analyses of the emotional impact of cognitive defusion techniques on positive and negative thoughts about the self. Part I of the thesis investigated the impact of defusion on explicit self-report measures. Experiment 1 examined the emotional impact of a brief defusion technique on positive and negative self-statements. In this case, defusion was manipulated through the use of pro- or anti-defusion instructions and prefixing the statements with defused or non-defused phrases (e.g. “*I am having the thought that I am a bad person*”). The results indicated that the defusion-related instructions had little or no impact on the explicit ratings, but the defusion prefix decreased statement discomfort while increasing statement believability and willingness. Indeed, this effect occurred only for negative statements, but positive statements remained unchanged on all measures. In Experiment 2, participants’ personalised self-statements were targeted for more direct intervention that involved defusion, thought control or placebo. In each case, the impact of instructions versus exercise was also assessed. In these findings, there was some superiority of defusion exercise over instructions, although defusion was generally associated again with decreased discomfort. However, willingness did not change and the believability of the statements now *decreased*. Thought Control overall decreased discomfort but had no impact on believability or willingness. No changes were associated with Placebo.

Part II of the current thesis attempted to examine the utility of defusion in the context of both explicit and implicit measures. Experiments 3 through 6 employed the Implicit Relational Assessment Procedure (IRAP) to determine the most suitable

types of self-relevant stimuli that could be targeted by the implicit measure while being susceptible to a defusion intervention. Participants' responses to positive and negative words were compared in the context of I AM versus I AM NOT (Experiment 3); I AM versus OTHERS ARE (Experiment 4); and I AM versus I SHOULD BE (Experiment 5). The results indicated strong positive implicit self-regard across all three contexts, although I AM versus I SHOULD BE showed the weakest levels of implicit positive self-regard, relative to the other two experiments. Experiment 6 replaced the target words with whole *statements* (e.g. 'I am so alone that it hurts') to determine if this would be associated with lower levels of implicit self-regard and indeed it was.

Experiment 7 provided a preliminary exploration of the susceptibility of the Statement-IRAP to defusion with groups of individuals with different levels of experience with defusion. Undergraduates (with no experience with defusion) were compared with a group of therapists (with considerable defusion experience). Although both groups reported strong implicit positive self-regard, those with more defusion experience were associated with stronger implicit outcomes. Experiment 8 investigated the impact of a defusion intervention or Placebo on implicit self-regard in individuals with no prior history of defusion. The data indicated that defusion increased implicit positive self-regard more than Placebo. The results presented here offer one of the first comprehensive experimental analyses of the impact of defusion on both explicit and implicit measures. Overall, the findings highlight the positive emotional benefits of defusion, relative to other more traditional coping strategies.

## Tables

		<b>Page</b>
Table 1	<i>Means (and Standard Deviations) for Psychological Measures for Each Instruction Group in Experiment 1</i>	51
Table 2	<i>Means (and Standard Deviations) for Discomfort, Believability and Willingness of the Three Presentation Formats for the Three Groups for Negative Self-Statements for Experiment 1</i>	53
Table 3	<i>Means (and Standard Deviations) for Discomfort, Believability and Willingness Ratings of the Three Presentation Formats for the Three Groups for Positive Self-Statements in Experiment 1</i>	56
Table 4	<i>An Outline of the Nine Experimental Conditions employed in Experiment 2</i>	70
Table 5	<i>Means (and Standard Deviations) for Psychological Measures for Each Condition in Experiment 2</i>	82
Table 6	<i>Means (and Standard Deviations) for Discomfort, Believability and Willingness Ratings at Pre- and Post- Intervention in Experiment 2</i>	84
Table 7	<i>The Stimulus Arrangements Employed in Experiment 3</i>	109
Table 8	<i>Mean Ratings (and Standard Deviations) for the Positive and Negative IRAP Words at Pre- and Post- IRAP in Experiment 3</i>	120
Table 9	<i>The Stimulus Arrangements Employed in Experiment 4</i>	124
Table 10	<i>Mean Ratings (and Standard Deviations) for Groups of Positive and Negative Words at Pre- and Post-IRAP in Experiment 4</i>	128
Table 11	<i>The Stimulus Arrangements Employed in Experiment 5</i>	132
Table 12	<i>Mean Ratings (and Standard Deviations) for Groups of Positive and Negative Words at Pre- and Post-IRAP in Experiment 5</i>	136
Table 13	<i>The Stimulus Arrangements Employed in the Statement-IRAP in Experiment 6</i>	146
Table 14	<i>Means (and Standard Deviations) for Each of the Psychological Measures for the Word-IRAP and Statement-IRAP Groups in Experiment 6</i>	151
Table 15	<i>Means (and Standard Deviations) for Each of the Psychological Measures for the ACT and Non-ACT Groups in Experiment 7</i>	161



Table 16	<i>Mean Scores (and Standard Deviations) for Believability Ratings of Positive and Negative Statements at Pre- and Post-IRAP for ACT and Non-ACT Groups in Experiment 7</i>	162
Table 17	<i>An Outline of the Ten Experimental Phases employed in Experiment 8</i>	172
Table 18	<i>Mean Scores (and Standard Deviations) for Each of the Psychological Measures for the Defusion and Placebo Conditions in Experiment 8</i>	175
Table 19	<i>Mean Scores (and Standard Deviations) for Believability Ratings of Positive and Negative Statements at Times 1, 2 and 3 for Defusion and Placebo Groups in Experiment 8</i>	176
Table 20	<i>Means (and Standard Deviations) for Discomfort, Believability and Willingness Ratings at Pre- and Post- Intervention in Experiment 8</i>	179

## Figures

	<b>Page</b>
<i>Figure 1</i>	An example of a negative self-statement in Abnormal presentation format in Experiment 1. 47
<i>Figure 2</i>	The experimental rating screen that followed the presentation of each automated self-statement in Experiment 1. 48
<i>Figure 3</i>	An example of the on-screen illustrations presented to participants in Experiment 3. 111
<i>Figure 4</i>	Examples of the four IRAP trial-types used in Experiment 3. 114
<i>Figure 5</i>	The mean $D_{\text{IRAP-(trial-type)}}$ scores and standard error bars for the sample stimuli in Experiment 3. 118
<i>Figure 6</i>	Examples of the four IRAP trial-types from Experiment 4. 125
<i>Figure 7</i>	The mean $D_{\text{IRAP-(trial-type)}}$ scores and standard error bars for the sample stimuli in Experiment 4. 126
<i>Figure 8</i>	Examples of the four IRAP trial-types used in Experiment 5. 133
<i>Figure 9</i>	The mean $D_{\text{IRAP-(trial-type)}}$ scores and standard error bars for the sample stimuli in Experiment 5. 135
<i>Figure 10</i>	Examples of the four Statement-IRAP trial-types used in Experiment 6. 148
<i>Figure 11</i>	The mean $D_{\text{IRAP-(trial-type)}}$ scores and standard error bars for the Word-IRAP and Statement-IRAP in Experiment 6. 150
<i>Figure 12</i>	The mean $D_{\text{IRAP-(trial-type)}}$ scores and standard error bars for the sample stimuli for the Non-ACT and ACT groups in Experiment 7. 159
<i>Figure 13</i>	The mean $D_{\text{IRAP-(trial-type)}}$ scores and standard error bars for the pre- and post- intervention IRAPs for the Defusion and Placebo groups in Experiment 8. 174

# **Chapter 1**

## **General Introduction**

# Chapter 1

## Introduction

When faced with troublesome or unpleasant thoughts or feelings, many individuals coerce themselves with instructions such as “Just don’t think about it” or “Concentrate on what’s positive”. But, despite the apparent intuitive sense in using strategies like suppression or distraction, few people really believe they can rid themselves of negative thoughts and feelings or can avoid them by simply thinking of something else. Even if such a strategy worked occasionally or in the short term, it seems unlikely that it is an effective way to manage psychological and emotional content. Indeed, there is a growing body of psychological evidence on *experiential avoidance* that suggests that this is the case. Specifically, recent clinical and experimental research suggests that: (1) experiential avoidance correlates with psychopathology; (2) avoidant strategies are unproductive and possibly counter-productive; (3) experiential acceptance, by contrast, is correlated with psychological health; and (4) cognitive defusion strategies are thought to play a crucial role in generating acceptance.

In line with this broad base of empirical evidence, the current thesis proposes three key arguments. First, experiential avoidance is correlated with poor mental health and as such does not appear to be a useful strategy for dealing with psychological problems. Second, acceptance (as an alternative to avoidance) is correlated with psychological well-being and reductions in psychopathology and thus appears to be a more effective strategy for dealing with psychological problems. Third, acceptance, whilst a commonly used term, has been only broadly defined and

the key active ingredients that mediate its outcomes have yet to be clarified. The thesis argues that cognitive and emotional defusion is one such ingredient that has not been clearly defined, but which does appear to be central to the positive effects of acceptance. In the sections that follow, the thesis presents these three core arguments and offers a review of the empirical evidence that supports these tenets.

### *Experiential Avoidance*

Experiential avoidance is said to occur when an individual is unwilling or unable to make psychological contact with private events (i.e. thoughts, feelings and emotions), particularly those that are negative and/or painful (Hayes, Strosahl, & Wilson, 1999). Avoidance is usually associated with attempts to change or eliminate these private events and the situations that cause them (Hayes, Wilson, Gifford, Follette, & Strosahl, 1996) and some common examples of avoidance that serve this purpose include: substance abuse, daydreaming, distraction and thought suppression.

Naturally, most individuals are more likely to engage in experiential avoidance with *negative* experiences that make them think or feel uncomfortable, compromised, worthless, futile, etc. than with *positive* experiences. Who wouldn't rather feel happy than sad? But this does not imply that being sad is a bad idea or that ignoring or removing thoughts and feelings of sadness will automatically make you feel happy. Indeed, there is growing empirical evidence that convincing yourself that you are not sad will not only *not* make you happy, but is more likely to exacerbate the sad feelings that you are experiencing (e.g. Purdon, 1999).

The basic idea behind this hypothesis is relatively simple and easily illustrated by a common example. Consider instructing a female friend as follows: "Do not think of cotton wool, think of anything else you like, but do *not* think of cotton wool". Of

course, her most likely reaction will be to automatically think of cotton wool and even if she managed to think of something else, she first has to think about cotton wool to then construct something that is not that. In other words, in order to successfully avoid a particular thought, your friend must generate a verbal rule that contains the to-be-avoided thought (e.g. “I must not think of cotton wool”). Furthermore, even if she tries to check that the target thought *has been* avoided, then she has to remember the rule and again the rule contains the to-be-avoided thought. So, when she remembers the rule, she ends up back where she started with the thought that she was trying to avoid in the first place.

Experiential avoidance appears to correlate with various forms of human psychopathology in both clinical and experimental contexts including: depression (Tull, & Gratz, 2008); Post-Traumatic Stress Disorder (PTSD – Orsillo, & Batten, 2005); and substance abuse (Sanchez-Craig, 1984). The potential role of experiential avoidance has also been highlighted in anxiety disorders (Salters-Pedneault, Tull, & Roemer, 2004), including the development and maintenance of Obsessive Compulsive Disorder (OCD – Shafran, & Rachman, 2004), agrophobia (Craske, Miller, Rotunda, & Barlow, 1990) and panic disorder (Tull, & Roemer, 2007).

A relationship has also been identified between the trauma associated with childhood sexual abuse and self-reported experiential avoidance (Marx, & Sloan, 2002). Specifically, avoiding trauma-related emotions is predictive of greater distress and severity of symptoms amongst individuals who have experienced a traumatic event (Gilboa-Schechtman, & Foa, 2001). A recent series of studies by Plumb, Orsillo and Luterek (2004) were concerned with the correlation between experiential avoidance and the psychological distress associated with traumatic events. In the first study of undergraduates, the researchers recorded a positive correlation between

propensity towards avoidance and level of psychological distress generated by a stressful life event (e.g. illness of a close family member). In the second study, they reported that the extent to which avoidance was employed as a strategy for coping with the trauma was a better predictor of subsequent distress than the severity of the traumatic event itself. Indeed, a subsequent study with war veterans confirmed a correlation between the use of experiential avoidance as a coping strategy for combat-related trauma and levels of depression and PTSD.

Evidence highlighting the negative impact of experiential avoidance also comes from research conducted in experimental contexts. For example, Karekla, Forsyth and Kelly (2004) investigated the impact of a carbon dioxide (CO<sub>2</sub>) challenge on a non-clinical sample of individuals rated as more or less inclined towards experiential avoidance or acceptance (as measured by the Acceptance and Action Questionnaire: AAQ; Hayes, Strosahl, et al., 2004). The results of the study indicated that those participants more pre-disposed towards avoidance than acceptance experienced more fear and panic-related symptoms and reported that they felt less in control than those more pre-disposed towards acceptance.

Other researchers have investigated the role of experiential avoidance with regard to sexual, physical and emotional abuse in a sample of inner-city substance users (Gratz, Bornovalova, Delany-Brumsey, Nick, & Lejuez, 2007). These researchers used both self-report and behavioural measures of avoidance -- the latter was designed to induce psychological and emotional distress using a modified version of the Paced Auditory Serial Addition Task (PASAT - originally developed by Gronwall, 1977). The results indicated that those individuals who had experienced moderate-to-severe sexual, physical and emotional abuse reported higher levels of experiential avoidance (as measured by willingness to persist on the PASAT),

compared to those reporting none-to-low abuse. Perhaps more importantly, other researchers have reported that experiential avoidance not only correlates with psychopathology, but also mediates the relationship between some forms of psychopathology and maladaptive coping styles. For example, the mediational role of avoidance has been highlighted in the context of anxiety severity (Kashdan, Barrios, Forsyth, & Steger, 2006), the relationship between observable anxiety reactions and depression (Tull, & Gratz, 2008), as well as the relationship between negative mood and PTSD in sexually assaulted women (Rosenthal, Cheavens, Lynch, & Follette, 2006).

Although avoidance behaviours (e.g. distraction, relaxation, substance abuse, or thought suppression) may be associated with the short-term alleviation of painful private events, evidence exists from the experimental, epidemiological and clinical literatures that avoidance behaviours are not associated with long-term, maintained psychological well-being (e.g. Hayes, Strosahl et al., 1999; Purdon, 1999). For example, although survivors of childhood sexual abuse rate avoidant coping strategies as the most effective in the short-term (Leitenberg, Greenwald, & Cado, 1992), these strategies remain associated with poor psychological functioning (Hayes et al., 1996). Furthermore, there is some empirical evidence that avoiding feared stimuli may paradoxically result in an *increase* of fear associated with the stimuli (Karekla et al., 2004), because avoidance may interfere with the physiological habituation to feared stimuli, thereby maintaining threatening meanings and interfering with new learning (Borkovec, & Hu, 1990). Indeed, engaging in avoidance can have a deleterious impact on long-term goals and valued behaviours (Wicksell, Renöfält, Olsson, Bond, & Melin, 2008). Taken together, therefore, experiential avoidance not only correlates with negative mental health outcomes and poor long-term psychological functioning,



but may also mediate these outcomes when it comprises an individual's coping strategy.

### *Thought Suppression as an Example of Avoidance*

Experiential avoidance can consist of any behaviour aimed at reducing, eliminating, or changing the form of private experiences and the situations that give rise to them (Hayes, Strosahl, et al., 1999). A number of topographies appear to characterise avoidance, although thought suppression is one of the most common (Zettle, 2007).

The limitations of using suppression to remove unwanted psychological content are well established (e.g. Wenzlaff, & Wegner, 2000). Specifically, attempts to suppress thoughts and feelings can result in a delayed increase in their occurrence, as well as an increase in their behavioural impact (Cioffi, & Holloway, 1993; Clark, Ball, & Pape, 1991). For example, in the classic 'white bear' experiment, Wegner, Schneider, Carter and White (1987) demonstrated that individuals who were explicitly instructed to suppress thoughts about a 'white bear' reported a paradoxical increase in these thoughts (both during, and immediately after, the suppression period – known as the '*rebound effect*') compared to those who had not received suppression instructions. Hence, suppression in this context was counterproductive because it actually *increased* contact with the unwanted content. Furthermore, similar effects have also been obtained with clinically relevant negative thoughts and moods (e.g. Wenzlaff, Wegner, & Roper, 1998).

Several researchers have attempted to investigate potentially broader correlations between thought suppression and a number of psychological disorders, including Generalised Anxiety Disorder (GAD) and OCD (Wenzlaff, & Wegner,

2000); PTSD (Ehlers, & Steil, 1995) and depression (Watkins, & Moulds, 2007). Beevers and Meyer (2004), for example, reported that high dispositional thought suppression (i.e. a strong propensity to suppress) combined with high life stress in a large sample of undergraduate students was associated with higher levels of depression. More disturbingly, other researchers have reported similar correlations between self-injurious thoughts and self-injurious behaviours amongst adolescents and young adults (Najmi, Wegner, & Nock, 2007). Specifically, in this study thought suppression appeared to mediate the relationship between emotional reactivity and the frequency of self-injury and suicidal ideation. Furthermore, adolescents who were more likely to suppress unwanted thoughts indicated that their self-injury functioned as escape from unwanted feelings, rather than attempts to communicate these feelings to others.

The negative effects of thought suppression have also been manipulated in experimental contexts (e.g. Wegner, Schreiber, Knutson, & McMahon, 1991). Research investigating physiological reactions (e.g. skin conductance levels) has highlighted negative effects associated with thought suppression strategies (Wegner, & Zanakos, 1994). For example, Campbell-Sills, Barlow, Brown and Hofmann (2006) investigated the effect of emotional suppression on self-report and physiological measures in a clinical sample of participants diagnosed with anxiety or mood disorders. The experiment simply involved instructing participants to apply a suppression-based coping strategy while watching an emotion-provoking film. The findings indicated that rather than diminishing distress during the film, suppression was associated with increased physiological arousal during the film, as well as persistent negative affect even after the film was over.

The mechanisms through which thought suppression works remain unclear. Although Wegner (1994) proposed the '*ironic processes theory*' which suggests two possible processes through which mental control works: (a) an operating process that searches for thoughts consistent with the desired state and (b) a monitoring process that searches for thoughts that are not consistent with the desired state, nevertheless, the precise conditions that best facilitate thought suppression have yet to be identified. Indeed, thought suppression efforts are not always associated with a paradoxical increase in thought frequency (see Abramowitz, Tolin, & Street, 2001 for a meta-analysis review). Some researchers have suggested that the effects of thought suppression may be influenced by the following variables: the context; the nature of the population (clinical or non-clinical); the nature of the to-be-suppressed thought (neutral or emotional -- see Purdon, 1999); mood (Wegner, Erber, & Zanakos, 1993; Wenzlaff, Wegner, & Klein, 1991); mental load (Wegner, & Erber, 1992); and motivation to suppress (Salkovskis, & Campbell, 1994). Furthermore, the 'rebound effect' depends upon the context in which the suppression occurred, because when an individual returns to this context the likelihood of re-experiencing the thought is high and the level of experiencing of the thought may be even higher than before (Roemer, & Borkovec, 1994; Wegner et al., 1991).

It also remains unclear whether suppression strategies simply *fail* to remove unwanted thoughts or whether they actually *increase* the rate or salience of the thoughts. For example, Purdon, Rowa and Antony (2005) investigated the impact of thought suppression on distressing intrusive thoughts reported by individuals with a primary diagnosis of OCD. The findings indicated that the frequency of unwanted thoughts did not actually increase for those participants instructed to suppress them. However, failure to successfully suppress the thoughts was associated with additional

distress about thought occurrences as well as a more negative mood state. Indeed, the researchers argued that because suppression efforts frequently fail, individuals become sensitised to their inability to control their thoughts and thus negative affect is created or exacerbated. In any case, the long-term negative impact of suppression efforts may not emerge until after the short-term utility of such strategies has passed.

Taken together, the negative outcomes for thought/emotion suppression strategies in both experimental and clinical contexts resemble, in no small way, the outcomes associated with experiential avoidance. Although the precise relationship between experiential avoidance and suppression is not well articulated in the literature, there appears to be considerable reason to believe that suppression very often functions as escape responding that promises relief from uncomfortable psychological content, but is in fact more likely to exacerbate it.

#### *Acceptance as an Alternative Coping Strategy*

The so called ‘third-wave’ of behavioural and cognitive therapies include Acceptance and Commitment Therapy (ACT; Hayes, Strosahl, et al., 1999); Dialectical Behavior Therapy (DBT; Linehan, 1993); Integrative Behavioural Couples Therapy (IBCT; Cordova, Jacobson, & Christensen, 1998); Mindfulness-based Cognitive Therapy (MBCT; Segal, Williams, & Teasdale, 2002); Functional Analytic Psychotherapy (FAP; Kohlenberg, & Tsai, 1991); as well as meta-cognitive approaches (Wells, 2000). Evidence across these approaches suggests that in promoting acceptance, rather than avoidance, of negative psychological content, positive outcomes for individual mental health are more likely.

According to ACT, avoidance is a natural outcome of our everyday verbal processes (Hayes, Strosahl, et al., 1999) that over time results in a narrow pattern of

behavioural responding. Despite the apparent intuitiveness and short-term benefits of avoidance strategies, ACT researchers have argued that these behaviours prevent individuals from engaging in more effective responses that produce life satisfaction (Hayes et al., 1996; Wicksell et al., 2008). Indeed, evidence from the clinical literature suggests that avoidance can complicate exposure-based strategies (Feldner, Zvolensky, Eifert, & Spira, 2003). Thus, clinicians and researchers have begun to investigate the utility of interventions aimed at decreasing experiential avoidance.

In ACT, therapists explicitly target experiential avoidance and suppression strategies as clinically relevant behaviour and attempt to orient clients towards acceptance as a more effective alternative in dealing with a whole range of psychological problems. Indeed, from an ACT perspective, a great many psychological problems may be summarised as manifestations of the generic condition of Experiential Avoidance Disorder (EAD: Luciano, & Hayes, 2001; Luciano, Rodríguez, & Gutiérrez, 2004) that can be described as an ineffective, over-generalised pattern of verbally regulated avoidance. In the language of ACT, clients are encouraged to accept, rather than avoid, their psychological content because the latter simply does not work. This switch in strategy then opens up the possibility that doing ‘something different’ *per se* (i.e. moving towards acceptance) will generate a broader and more flexible repertoire of responding that will increase the possibility for behavioural change.

There is considerable reason to believe that acceptance, in and of itself, is correlated with positive outcomes in terms of dealing effectively with psychological content. For example, in the context of pain, acceptance is associated with lower pain-related avoidance, anxiety, depression, disability and pain intensity, as well as a greater likelihood of, and faster, return to work (McCracken, 1998; McCracken, &

Eccleston, 2003). Acceptance has also been associated with a lower probability of having a psychiatric disorder (Donaldson-Feilder, & Bond, 2004). Furthermore, higher levels of acceptance have been associated with lower levels of depression (e.g. Polusny, Rosenthal, Aban, & Follette, 2004) and anxiety (e.g. Stewart, Zvolensky, & Eifert, 2002).

In a non-clinical application of ACT, Bond and Bunce (2000) investigated the impact of an ACT protocol on coping with stress in workplace environments and reported significant improvements in participants' mental and general health over a three-month period. Furthermore, the researchers reported that the positive outcomes were mediated by the acceptance of undesirable thoughts and feelings, and not by efforts to control the occurrence of such thoughts. Marcks and Woods (2005) also demonstrated the positive effects of acceptance within an experimental context when they compared an acceptance versus suppression strategy for coping with personal intrusive thoughts in a non-clinical sample of undergraduate students. Participants' single most upsetting intrusive thought was identified using the Revised Obsessive Instructions Inventory (ROII; Purdon, & Clark, 1994). Study 1 investigated the relationship between naturally occurring coping strategies and intrusive thoughts. A higher disposition towards acceptance was associated with fewer obsessional symptoms, lower levels of depression and anxiety, and lower discomfort with respect to the intrusive thought compared to participants who naturally suppress these thoughts. In Study 2, the same students were randomly assigned to one of three conditions: acceptance, suppression, or monitor-only. As expected, explicit instructions to accept intrusive thoughts was associated with a decrease in the discomfort associated with the thoughts, but not in their frequency.

As a fuller treatment regime, ACT has proven equally effective in clinical contexts involving depression (Zettle, & Hayes, 1986); chronic pain (Dahl, Wilson, & Nilsson, 2004; McCracken, Mackichan, & Eccleston, 2007); substance abuse (Hayes, Wilson, et al., 2004); stress (e.g. Bond, & Bunce, 2000); psychotic symptoms (Bach, & Hayes, 2002); trichotillomania (Woods, Wetterneck, & Flessner, 2006); and maths anxiety (Zettle, 2003). ACT interventions have also proven to be effective in the decrease of long-term disability associated with drug refractory epileptic seizures (Lundgren, Dahl, Melin, & Kies, 2006), smoking cessation (Gifford et al., 2004) and improving diabetes self-management (Gregg, Callaghan, Hayes, & Glenn-Lawson, 2007).

For example, in the treatment of trichotillomania, Woods et al. (2006) compared the impact of ACT/plus habit reversal training with a waitlist control group. Participants in the intervention condition reported a significant decrease in hair pulling that coincided with reductions in experiential avoidance, anxiety and depression. Furthermore, these gains were generally maintained at a three-month follow up. The positive results reported for ACT in the treatment of in-patients with psychotic symptoms are particularly impressive. Specifically, Gaudio and Herbert (2006) compared ACT/plus enhanced treatment as usual (ACT/ETAU) with ETAU alone. The former group showed greater overall symptom improvement, including lower distress for hallucinations and lower re-hospitalisation rates following a four-month period. Taken together, these and the other cited studies suggest the utility of acceptance-based strategies, especially when contained within an ACT treatment program, for reducing symptoms associated with psychopathology. Although the experimental and process-based evidence suggests that at least some of the improvements are mediated through acceptance, the processes that underpin

acceptance remain unclear. In the sections below, the thesis presents the argument that the concept known as *cognitive defusion* forms a core feature of the acceptance of psychological content.

### *Cognitive Defusion versus Cognitive Fusion*

Cognitive defusion strategies are a critical feature of ACT and are designed to facilitate acceptance, particularly of negative psychological content. Defusion strategies encourage clients to change the context or perspective from which their negative thoughts are observed, rather than changing the form or frequency of the thoughts directly. In this way, defusion strategies are unlike avoidance strategies, and resemble much more the type of approach behaviour encouraged in traditional exposure therapies. Defusion is designed to help individuals see private events for what they are (e.g. seeing a thought as just a thought), rather than as something that must be avoided. Thus, in some ways, defusion strategies resemble techniques used in cognitive-based and mindfulness-based therapies. For example, cognitive therapists use distancing strategies that involve ‘*stepping back*’ from dysfunctional thoughts in an attempt to see them as thoughts, rather than as facts (e.g. Beck, 1970; Kross, Ayduk, & Mischel, 2005). Furthermore, therapists working from within the mindfulness framework try to establish a decentered perspective that facilitates metacognitive awareness (Teasdale et al., 2000) and encourages present moment awareness of external and internal stimuli (Kabat-Zinn, 1994).

According to ACT, defusion is particularly important in dealing with painful or traumatic events because it creates a safe and workable perspective from which the psychological content can be observed as *separate from the observer*. For example, in ACT, the phrase “you are more than what your thoughts say you are” is often



employed. Put simply, just as the whole is greater than the sum of the parts, a whole human being is greater than all of the thoughts and feelings s/he may have. Indeed, in cases where psychological content is particularly traumatic, it is frequently necessary to create a dissociation between the thinker and the thoughts before acceptance of the thoughts can be achieved. In order to do so, ACT therapists often rely on the use of metaphors. For example, 'The House and the Furniture' metaphor is commonly used by ACT therapists to highlight the distinction between individuals and their thoughts:

It's as if you are a house, filled with furniture. The furniture is not, and can never be the house. Furniture is the content of the house and the house merely holds or contains it. It provides the context in which the furniture can be furniture. Whether the furniture is thought to be good or bad, says nothing about the value of the house. You are the house but not the furniture. Your thoughts and feelings are the furniture. Just as the furniture is not the house, your thoughts and feelings are not you. They are simply experiences you have that are like pieces of furniture.

Once a sense of separation between thinkers and thoughts has been created, ACT therapists attempt to weaken these associations further by explicitly employing cognitive defusion techniques.

The concept of *cognitive fusion* is used to describe situations in which an individual's overt behaviours are under the direct control of private internal events (Hayes, Strosahl, et al., 1999). From an ACT perspective, negative thoughts in and of themselves are not problematic, but difficulties arise when thoughts are believed to be true indicators of the self (i.e. when clients are *fused*). In this sense, fusion is similar to the idea of 'thought-action' fusion (TAF: Shafran, Thordarson, & Rachman, 1996) first proposed within the context of clinical obsessions. According to this perspective, TAF is based on the incorrect assumption that: (1) thoughts can directly influence external events and (2) experiencing negative intrusive thoughts is morally equivalent to engaging in that behaviour. Research has reported a small to moderate relationship between TAF and OCD (Rassin, Merckelbach, Muris, & Schmidt, 2001); anxiety

disorders (Rassin, Diepstraten, Merkelbach, & Muris, 2001); and depression (Abramowitz, Whiteside, Lynam, & Kalsy, 2003).

From an ACT perspective, fusion also becomes problematic when actions are appraised as evidence of reality, thus giving indirect, but equal, credence to the thoughts that generated the actions in the first place. Hence, when individuals identify themselves as co-ordinated with their own behaviour (e.g. 'I am what I do'), then indirectly they are also co-ordinated with the psychological content that directed it. As a result, the self may become fused with thoughts and an individual might come to believe that 'I am what I think'. Naturally, the psychological and emotional outcomes associated with this type of thinking can be devastating (e.g. "I think I can't cope with my children, therefore I must be a terrible mother"). Indeed from an ACT perspective, cognitive fusion is considered to play a causal role in avoidance behaviours and as such warrants direct clinical intervention (Orsillo, Roemer, Block Lerner, & Tull, 2004).

Although there is some clinical evidence to suggest the utility of defusion and similar strategies in promoting psychological well-being, there remains no clear evidence of how these strategies work in the context of therapy. Specifically, although there are a number of theories that have attempted to explain how defusion strategies work, the detail and supporting evidence remain exploratory (for a review see Blackledge, 2007).

### *Fusion as a Language Process*

The concept of cognitive fusion, as defined by ACT, is not inconsistent with the modern theory of language and cognition known as Relational Frame Theory (RFT; Hayes, Barnes-Holmes, & Roche, 2001). According to RFT, human language

is inherently relational and incorporates relations of difference, co-ordination, opposition, comparison, perspective-taking and so on. From an RFT perspective, the fact that any stimulus can become related to another stimulus accounts for the highly arbitrary and generative nature of human language. For example, we can use language to discuss either previous events that we have already experienced or future events of which we have no direct experience.

A critical feature of the RFT account of verbal relations involves the transformation of stimulus functions, in which functions associated with one stimulus are readily transferred or transformed to another stimulus when the two are related (even arbitrarily). These transformations may be the process that underlies cognitive and emotional fusion. Consider the following scenario:

Imagine a married woman with children who does not work and who begins to have rushing thoughts one day while doing the grocery shopping (e.g. “I don’t see why I have to do all the menial family tasks like shopping. I could do more with my life than this”). Even though the woman may have had these thoughts before, they somehow on this occasion feel particularly salient (e.g. perhaps she had tense words with her husband the night before on an unrelated matter). When the woman returns home, she begins to feel troubled by her thoughts and starts to feel tension across her chest. When she thinks through where the thoughts are coming from, she begins to piece together perhaps unrelated events, including: tense words with her husband the night before; the fact that he had been away with friends without her several weekends ago; and his recent stress at work which has meant that he has had little interactions with the children. As these recollections coincide, the woman begins to feel even more tense and new thoughts such as “Perhaps my husband is having an affair” or “Perhaps I don’t love him anymore” also occur. Naturally, these latter thoughts strike the woman as more serious and thus she becomes more tense and increasingly concerned that “Perhaps I really don’t love him anymore. What will I do?” Now consider when the husband returns home from work that evening and the woman finds it difficult to address him because she feels guilty about the thoughts she had that day regarding their relationship. So, she remains somewhat aloof and the husband, stressed out from another day at work, retorts about the woman’s mood and the fact that he feels it is misplaced because all she had to do all day was the shopping. Under these circumstances, the woman might feel deeply hurt by this statement particularly because it resonates with her own feelings of futility in the grocery store. Rather than retort, however, because she is afraid to give away clues to what she is really thinking, the woman thinks “Someone who loved me would not say that I was useless” and so the cycle of marital uncertainty and distress deepens.

In the example above, there are many loose connections or relations between thoughts, feelings, actions and other events in the world. For example, it just happened that the woman felt pangs of tension in the grocery store, if she had been in the post office, this may not have happened. What is most important is that in her

mind the woman connected the feelings with the grocery store and began to try to make sense of the connection between these two stimuli. These derived relations then generated new connections with new meanings that then gave even more meaning to the original relations. From an RFT perspective, this added meaning is unavoidable by virtue of the fact that the events are related. In short, once the relations happen, meaning in one event automatically changes meaning in another. So, very quickly, the initially arbitrary relations now appear to have great meaning and convince the unsuspecting believer that related thoughts are true (e.g. “if I have so many thoughts and if they make me feel tense, they must be true”). Perhaps more distressingly, the relations and meanings are now interpreted as saying something important about who the woman is (“I don’t love my husband anymore and he probably doesn’t love me”). As a result, the thoughts may become even more strongly connected to overt action. For example, imagine that the woman then discusses this matter with an unmarried friend who confirms that the marriage comprised of a poor match and that it was the case that the relationship should end.

For RFT, there are important and unavoidable relationships among thoughts/feelings, overt action and the self that appear to underpin cognitive and emotional fusion. Consider the relationship between thoughts and feelings. If you have one thought and you feel bad and you have a second thought that participates in a relation of co-ordination with the first thought, then the aversive functions of the first thought (feeling bad) immediately transfer to the second thought and now *it* also makes you feel bad. In our evolutionary history, it makes sense for humans to avoid aversive events. For example, if we feel too hot or too cold, or hungry etc., we work quickly to get the situation resolved – and this is likely to benefit the organism in the short and long terms. But applying the same strategy to verbal threats (that have no

actual basis in reality) is a type of *over-generalisation* that makes much less evolutionary sense. Furthermore, as highlighted in the former section on avoidance, attempts to use such strategies do not work for psychological content anyway. But, if you then engage in experiential avoidance (e.g. suppression) to remove feeling bad and the thoughts that go with it, especially when it appears to work in the short term, then you are likely to use the same strategy to remove related thoughts that also make you feel uncomfortable. Hence, the relationship between internal content and action is strengthened, even if it doesn't work particularly well. For RFT and ACT therefore, a critical problem in psychopathology is the fusion amongst thoughts, action and the self, and our attempts to avoid these relationships and events in futile efforts to make them go away.

According to RFT and ACT, fusion is a natural feature of verbally sophisticated language abilities and is almost unavoidable. Indeed, the high prevalence of psychological problems in the general population supports this view. For example, almost half of all Americans (46.4%) will meet the criteria for a psychological disorder sometime during their lives as identified by the Diagnostic and Statistical Manual (DSM-IV; American Psychiatric Association, 1994) with onset usually appearing during childhood or adolescence (Kessler, Berglund, Demler, Jin, & Walters, 2005). However, not all of our negative thoughts start us out on a cycle of critical uncertainty, in part because: they are not associated with salient emotions; they have only limited connections with actions; and when we express them openly to others they are less likely to be suppressed. Acceptance, therefore, is an important prophylactic against fusion because when we accept the arbitrary connections between our thoughts and emotional states, they are less likely to control our behaviour. Defusion strategies are thought to work by highlighting the common

process of verbally relating events to other events (e.g. relating thoughts to the self-concept) and thereby disrupting the existing problematic verbal relationships (see Hayes, Strosahl, et al., 1999). Put simply, as soon as we face our thoughts and feelings and see them for what they are, we realise that they are meaningless unless we try to attach some meaning to them.

### *Examples of Defusion Exercises*

Acceptance and Commitment Therapy incorporates a range of cognitive defusion exercises, including cognitive distancing and word repetition, to target problematic areas of fusion. All defusion techniques are focused initially on altering clients' perspectives with regard to their thoughts and feelings, from a position in which they *are* them (fusion) to a position in which they *have* them (defusion) – remember the metaphor of the house and the furniture.

When this small distance has been placed between the thoughts and the thinker (but not from the perspective of pretending that they are not related), then acceptance of even traumatic content becomes more likely. In order to begin to establish this new perspective, ACT therapists encourage clients to see all thoughts as thoughts and all feelings as feelings (Hayes et al., 2001). For example, clients are often encouraged to prefix descriptions of thoughts with the phrase: “I am having the thought that...”. So, if a client is struggling with the recurrent thought: “I am a useless human being”, she will be encouraged to rephrase this statement as: “I am having the thought that I am a useless human being”. This transforms the fused functions of the thought in a way that alters the relation between the thought and the self from one of co-ordination to one of distinction. For example, previously the client was interpreting the relationship between the house and the furniture as the same, and not appreciating that one

actually *contains* the other. So, the therapist tries to alter the relationship from one that is co-ordinated to one that is most likely hierarchical. Hence, from this new perspective the client may derive “I am actually bigger than my thoughts, they need not be everything that directs my actions”.

More extensive experiential exercises are also employed in ACT to highlight how easily our internal content (often labelled as our ‘mind’) gets the upper hand over our actions in a relentless series of directions. For example, clients frequently report racing and unco-ordinated thinking and ACT therapists use exercises such as ‘*taking your mind for a walk*’ to indicate the relentless continuity of the mind’s instructions and evaluations (Hayes, Strosahl, et al., 1999). Specifically, this exercise requires the client to walk around the room with the therapist following behind and pretending to be the client’s mind thinking aloud. As the therapist says all the thoughts aloud, the client is instructed to simply listen to the ‘mind’ and to continue to do what she was doing (e.g. continue walking around the room, even if the ‘mind’ tells you to sit down). Other such exercises include imagining that thoughts are like leaves floating down a stream (Hayes et al., 2001) or soldiers marching in a parade (Hayes, Strosahl, et al., 1999) – all designed to establish a broader sense of self, within which private events have less importance. Cognitive defusion techniques, therefore, are inherently acceptance-based and thus depart from traditional therapeutic methods that encourage clients to change the form or frequency of their thoughts, eliminate them altogether, or replace them with more likeable content.

Word repetition exercises are also employed by ACT therapists to facilitate defusion. Such exercises are based on Titchener’s (1916) rapid word repetition technique, also known as *semantic satiation*. Semantic satiation has been defined as a repetition induced loss or reduction in the meaning associated with words, which in

turn is thought to reduce access to related semantic information (Cohene, Smith, & Klein, 1978; Smith, 1984). Thus, it is possible that even a familiar word can become meaningless through weakening the associative context of that word. Although Titchener's (1916) repetition technique was originally designed to investigate illusions of recognition and memory with neutral stimuli, the technique has been adapted for therapeutic purposes. For example, Benson and Klipper (1976) highlighted the benefits of using single word mantras to focus one's attention, for the purpose of relaxation. Indeed, single word mantra techniques have been successfully applied within clinical contexts. For example, in a sample of patients with heart disease, the frequency of premature ventricular contractions was reduced following training in the relaxation response using single-word mantras (Benson, Alexander, & Feldman, 1975).

From an ACT perspective, word repetition exercises are useful for facilitating defusion because the meanings and functions of the word reduce when the word becomes a sound (Hayes, Strosahl, et al., 1999). For example, clients may initially explore all the properties associated with the word "milk" (e.g. cold, creamy, frothy, etc.) followed by a short period of rapidly repeating the word "milk" aloud until it loses all meaning. The therapist then uses this experience to emphasise that the client's negative content is purely verbal and not a reflection of reality. Thus, from an ACT perspective, repetition exercises are not employed to reduce the distress associated with psychological content, nor are they an attempt to remove the client's unwanted content. Indeed, research has indicated that the effects of semantic satiation are temporary and are seen to wear-off after time (e.g. Esposito, & Phelton, 1971). Rather, the person makes contact with only the direct functions of the word (i.e. sees the word as simply a word), thus the derived relational functions of the word are, at



least temporarily, not available. The function of the word/thought changes without changing the form of the thought, or reducing its frequency. It is believed that this switch in strategy allows the person to see thoughts for what they are, rather than as something that must control behaviour, thereby increasing the behavioural options available to the individual.

### *Empirical Evidence of Cognitive Defusion*

Although cognitive defusion strategies are frequently used within an ACT-based clinical context, the empirical evidence to support their efficacy is relatively limited. To date, only a number of research studies have attempted to investigate the processes of cognitive defusion. *Indirect* evidence of defusion exists in the form of ACT-based clinical outcome trials (e.g. Bach, & Hayes, 2002) and clinical analogue studies (e.g. Gutierrez, Luciano, Rodriguez, & Fink, 2004). For example, chronic pain analogue studies involving the Cold Pressor Task have indicated that acceptance-based interventions that rely heavily on defusion increase participants' pain tolerance (Hayes, Bisset et al., 1999; Takahashi, Muto, Tada, & Sugiyama, 2002).

Bach and Hayes (2002) investigated the impact of a brief acceptance intervention on the rate of rehospitalisation of psychotic in-patients. In this study, treatment as usual (TAU) was compared to TAU plus an ACT intervention that relied heavily on cognitive defusion (e.g. clients engaged in the *take your mind for a walk* exercise). The results indicated that participants assigned to the ACT condition showed a 50% reduction in the rate of rehospitalisation compared to TAU alone during a 120-day follow-up. The researchers attributed this to an increase in levels of acceptance and a decrease in levels of believability of psychotic symptoms. In other words, the therapy was not explicitly trying to decrease the number of psychotic

thoughts, but was trying to reduce the verbal hold these had over the individual and his/her actions. Of course however, reducing the extent to which thoughts are believed will almost by definition reduce the distress associated with the thoughts in the first place. The research, not surprisingly, also recorded a decrease in distress associated with ACT, even though distress was not targeted directly.

Taken together, these studies suggest that ACT interventions that rely heavily on defusion result in a decrease in levels of believability and/or discomfort associated with psychological content. However, in this type of research, defusion exercises were presented as one aspect of an overall ACT package, thus there remains limited evidence for the utility of defusion as an isolated component.

More direct evidence for defusion was generated by the first empirical investigation of defusion strategies when Titchener's (1916) word repetition exercise was used in the context of negatively evaluated self-referential content (Masuda, Hayes, Sackett, & Twohig, 2004). Specifically, undergraduate students were asked to generate two self-relevant negative thoughts that they found particularly disturbing (e.g. "I am too fat") and then to restate each thought in a single word (e.g. "fat"). Participants then rated the levels of discomfort and believability associated with each word. In Experiment 1, the researchers compared the defusion strategy (i.e. defusion rationale combined with the 'milk, milk, milk' word repetition strategy) with a distraction technique (i.e. reading an unrelated article about Japan). In Experiment 2 the same defusion strategy was compared with a thought control strategy (e.g. participants could use a variety of strategies including positive imagery, breathing exercises, or positive self-talk). Both experiments employed an alternating treatment design, with four participants. One of the two self-generated negative thoughts was randomly assigned to each treatment condition. In both experiments, participants were

exposed to three presentations of each intervention in a random sequence, but with no more than two consecutive exposures to any one strategy. The results from both experiments indicated that the defusion rationale produced the largest reductions in the believability of the negative self-relevant words, as well as in the levels of discomfort produced by the words.

### *The Current Thesis*

Although research has begun to investigate cognitive defusion as an isolated component and results suggest that defusion is a psychologically active construct (Masuda et al., 2004), there remains limited knowledge about the processes through which it works. The focus of the current thesis, therefore, was to determine the conditions under which defusion works best and to understand the underlying processes associated with it. To this extent, Part I of the present research investigated the self-reported emotional impact of cognitive defusion. Experiment 1 had three main aims: (1) to investigate the impact of an alternative defusion strategy to that employed by Masuda et al. to determine if similar effects would be obtained across the various defusion exercises; (2) to investigate the impact of simple defusion-related instructions; and (3) to investigate the impact of defusion on positive, as well as negative, self-statements. Similar to the results obtained by Masuda et al., the defusion strategy in this first study also resulted in a decrease in the emotional impact of the negative self-statements. However, the defusion-related instructions had no impact on the overall ratings. Possible shortcomings in the first study included the possibility that both the instructions and the intervention were too brief, and that the impact of defusion may have been greater if *personally relevant* psychological content was targeted.

Experiment 2 attempted to address these issues by: (1) investigating the impact of defusion on personally relevant negative thoughts; and (2) systematically examining the relative utility of defusion instructions versus defusion exercises. In addition, Experiment 2 compared defusion to other strategies (e.g. thought control). With these modifications in place, defusion was again associated with a decrease in the emotional impact of negative self-referents. Furthermore, a slight superiority of defusion exercises over defusion instructions was recorded. Defusion also proved to be superior to the other coping strategies targeted.

One issue that was consistent across both Experiments 1 and 2 was the use of explicit measures to assess participants' reactions to their psychological content. Although these tools are commonly employed in clinical research and practice, it may be reasonably argued that the self-report measures used to assess the impact of defusion are susceptible to the same generic weaknesses that potentially influence all explicit methodologies. Part II of the current thesis, therefore, investigated the use of alternative measures of believability, distress, etc. that can be used as indices of the impact of cognitive defusion strategies. Specifically, Experiments 3 to 8 (presented in Chapters 4, 5 and 6) utilised the Implicit Relational Assessment Procedure, an implicit measure of cognition to measure defusion-related change (IRAP; Barnes-Holmes et al., 2006). Further details of using implicit measures of cognition as an alternative to explicit measures will be presented in Chapter 4.

Experiments 3 through 6 were designed to determine which self-relevant stimuli presented in different contexts would potentially be most relevant for use with a defusion manipulation. Results from these experiments suggested that an IRAP presenting positive and negative self-statements in the context of I AM and I SHOULD BE stimuli would be most susceptible to a defusion manipulation. Based on

these results, Experiment 7 (Chapter 5) was designed as an exploratory experiment to determine the potential susceptibility of implicit measures to cognitive defusion. Specifically, the experiment compared two groups of individuals that differed with regard to their experience with cognitive defusion. The ACT Group (thought to have considerable experience with defusion) reported stronger implicit positive self-regard than the Non-ACT Group (considered to have no experience with defusion). Neither group produced strong reactions to how they felt they should be. Finally, Experiment 8 investigated the impact of defusion on implicit self-regard in individuals with no experience with defusion. Specifically, the experiment compared Pre- and Post-Intervention responses to Defusion or Placebo conditions. Results indicated that defusion was associated with a larger increase in implicit positive self-regard and a smaller increase in implicit attitudes that participants should be positive relative to placebo. The seventh and final chapter summarises and discusses the findings from all eight experiments and suggests directions for future research in this area.

# **PART I**

## **Chapter 2**

### **Experiment 1**

**An Experimental Test of a  
Cognitive Defusion Exercise:  
Coping with Positive and Negative  
Self-Statements**

## Chapter 2

### Experiment 1

# An Experimental Test of a Cognitive Defusion Exercise: Coping with Positive and Negative Self-Statements

## INTRODUCTION

Only one published study has investigated the specific impact of cognitive defusion on negative psychological content (Masuda et al., 2004). The results of this research were consistent with predictions made by ACT and RFT and indicated that defusion *decreased* the emotional impact of negative self-relevant thoughts in terms of baseline levels of discomfort and believability associated with the thoughts. In order to manipulate defusion, the researchers employed Titchener's (1916) rapid word repetition technique (e.g. milk, milk, milk), originally designed to reduce the semantic functions of words and commonly used in ACT as a defusion technique.

Word repetition exercises are only one of a number of defusion exercises designed to create cognitive distance between individuals and their psychological content. Hence, it remained possible that the results reported by Masuda et al. (2004) were specific to word repetition, rather than to defusion in general. Put simply, different outcomes may be associated with different types of defusion exercises. With this in mind and the fact that there is limited existing evidence of the *utility* of defusion, Experiment 1 employed an alternative defusion strategy.

Rather than using word repetition, the current study employed an alternative defusion technique often employed by ACT therapists. Specifically within ACT, the



prefix phrase ‘*I am having the thought that*’ is used to create defusion of problematic psychological content. Similar to word repetition, the defusion prefix is used to highlight that the subsequent statement is simply a thought and not a fact.

### *The Current Study*

Experiment 1 was designed to test the emotional impact of the defusion prefix (‘*I am having the thought that*’) on a set of generic self-statements. The study had three main aims: (1) to determine if the results reported by Masuda et al. (2004) were specific to word repetition techniques or were outcomes that could be attributed to a range of defusion techniques; and (2) to systematically compare the utility of defusion instructions versus a defusion exercise. In ACT in general, defusion is delivered by means of experiential exercises, rather than by instruction and the results of the Masuda study indicated that defusion instructions alone were of limited utility. Nonetheless, it remained possible that positive defusion outcomes may be more easily achieved by simply instructing participants on defusion, rather than employing defusion exercises. (3) The current study also attempted to determine the relative impact of defusion on positive versus negative self-statements. In this way, the positive statements functioned as a type of experimental control to determine whether defusion affects all, or only specific, types of psychological content.

The current study involved the presentation of a series of self-statements on a computer screen. In one condition (referred to as *Normal* presentation) the self-statements appeared in their original form (e.g. “I am a bad person”). In the *Defused* presentation, the statements were altered by attaching the defusion prefix “I am having the thought that” (e.g. “I am having the thought that I am a bad person”). However, because the defusion prefix changed the length of the statement and might

also be seen as creating an unusual verbalisation, a third condition (*Abnormal* presentation) was created in which the self-statements were prefixed again, but with a phrase that matched the defusion statement in length and that was equally unusual (i.e. “I have a wooden chair and I am a bad person”). This acted as an experimental control by providing a comparison statement of similar length that was also unusual, but did not possess the same level of defusion as the defusion prefix. We predicted that the Defused presentation would have a greater impact on the emotional ratings associated with the target self-statements than the other two types of presentation.

In order to control for demand compliance effects (e.g. Fernandez, & Turk, 1994; Kanter, Kohlenberg, & Loftus, 2004), participants in the current study were also divided according to the types of explicit instructions they received regarding the utility of the defusion technique. Each instruction stated that previous research had shown that the defusion prefix: (1) increased (referred to as *Pro-Defusion* Group); (2) decreased (*Anti-Defusion* Group); or (3) had no effect upon (*Neutral* Group) the emotional impact of self-statements. We hypothesised that the defusion prefix (considered here as a defusion exercise) would out-perform defusion instructions in terms of the emotional impact of the self-statements.

Traditionally, within a therapeutic context, cognitive defusion exercises have been employed as coping strategies for *negative* thoughts and feelings. Thus, we made no clear predictions about the impact of defusion on the positive self-statements. These statements were simply included as an experimental control to determine if the emotional impact of defusion was dependent upon psychologically engaging with negative thoughts and feelings, or simply reflected the process of distancing oneself from one’s thoughts in general. Overall, the inclusion of the

positive self-statements would allow a fuller understanding of the manner in which defusion works.

The dependent variables employed in the current study involved measuring participants' reactions to the self-statements, using three self-report Likert-type rating scales that assessed associated levels of *discomfort*, *believability* and *willingness* (Hayes, Bisset et al., 1999; Levitt, Brown, Orsillo, & Barlow, 2004; Masuda et al., 2004). We hypothesised that the Defused presentation would differentially impact the explicit ratings of the negative statements, relative to the Normal and Abnormal presentations. Specifically, we hypothesised that the negative statements in the Defused presentation would likely decrease believability to see or say the statements, but would increase willingness, relative to Normal and Abnormal presentations. The believability prediction was consistent with ACT, in which changing one's perspective with regard to one's psychological content by indicating that the person is greater than the sum of the content, reduces the believability of the content (Hayes, Luoma, Bond, Masuda, & Lillis, 2006). In other words, if I know that the thought "I am a horrible person" is only a thought, then I will be aware that the thought can say nothing about who I am as a real person. As a result, the thought will be less believable. Put simply, the believability of thoughts results from their potential to be viewed as having more meaning than they need have (i.e. by saying something about who you are as a person). When this potential is reduced in the context of particular psychological content, then that content, by definition, will be less believable. In the research by Masuda et al. (2004), believability of the psychological content that was targeted was reduced as a result of the defusion exercise.

We also hypothesised that the defusion prefix would increase participants' willingness to make psychological contact with the negative statements. This

willingness prediction was also consistent with ACT and simply asserted that if you are defused from psychological content such that it says nothing about who you are, then there is no need for you to be unwilling to have it. In other words, if the content has no power, then you need not be afraid of what it can do to you. Although, willingness had not been targeted directly by Masuda et al. (2004), other sources of empirical evidence indicated that acceptance-based strategies increase willingness (Levitt et al., 2004).

In the current research, we initially made no clear predictions regarding the impact of defusion on psychological discomfort. From an ACT perspective, discomfort is not targeted directly because one can still engage in valued action when levels of discomfort are high. Furthermore, because individuals generally have little or no control over their emotional states, ACT encourages clients not to focus upon them as a potential source of behaviour change. According to this view then, discomfort may increase, decrease, or remain unchanged. Specifically, one might argue that increasing acceptance of content might *increase* discomfort because one is more willing to make experiential contact with it. Alternatively, one might predict that discomfort would *decrease* when one realises that thoughts associated with emotion are only thoughts, and as such have limited control or impact on who you really are. Interestingly, Masuda et al. (2004) reported that their defusion technique resulted in decreased discomfort. Taken together then, we had no clear initial prediction about what would happen to participants' levels of discomfort as a result of the defusion technique.

## METHOD

### *Participants*

Sixty participants (31 female, 29 male) aged between 18 and 57 years old comprised the final sample (mode = 21). All were undergraduate students, recruited through Faculty announcements in the Department of Psychology at the National University of Ireland Maynooth (NUIM). None had prior exposure to similar experimental procedures. Twenty participants were each assigned to three groups that differed only in terms of the defusion-relevant instructions provided to them (Pro-Defusion, Anti-Defusion and Neutral). No incentives were offered in return for participation. In addition to the 60 participants who completed the experiment, four participants indicated that they did not understand the defusion-related instructions provided at the beginning of the study, thus were not included in the final data set.

### *Setting*

Experiment 1 was conducted in an experimental room that was in a quiet location and generally free from distraction. All participants completed each aspect of the study independently. The experimenter remained seated outside the room at all times. All participants were seated on a standard plastic office swivel chair (rather than a wooden chair), because of the wording of the Abnormal presentation (see below).

### *Apparatus and Materials*

*Psychological Measures.* Each participant first completed three self-report questionnaires. These comprised of the Acceptance and Action Questionnaire (AAQ-

35; Hayes, Strosahl, et al., 2004 -- Appendix A), the Beck Depression Inventory (BDI-II; Beck, Steer, & Brown, 1996 -- Appendix B) and the State Trait Anxiety Inventory (STAI Form Y-2; Spielberger, Gorsuch, Lushene, Vagg, & Jacobs, 1983 -- Appendix C). The assignment of participants to one of the three experimental groups did not depend upon their questionnaire scores. The measures were simply included to control for any variability across groups that might account for any differences obtained during subsequent experimental phases.

The *AAQ-35* is a self-report measure of an individual's general level of emotional avoidance. Individuals are asked to rate the truth of each of 35 statements as it applies to themselves, on a scale of 1 (never true) to 7 (always true). The scales contain 17 negatively reversed items. Total scores range from 37 to 245, with low scores indicating high avoidance and low acceptance, and high scores indicating low avoidance and high acceptance. Although the *AAQ-35* is a relatively new measure, initial research on similar versions of the measure indicate good psychometric properties, with good evidence of convergent, criterion-related and construct validity (see Bond, & Bunce, 2003; Hayes, Strosahl, et al., 2004).

The *BDI-II* consists of 21 groups of four statements (all scored from 0-3). Individuals are asked to select one statement from each group that best describes how they have been feeling during the previous two weeks. The *BDI-II* is scored according to four categories: 0-13: Minimal Depression; 14-19: Mild; 20-28: Moderate; and 29-63: Severe. The *BDI-II* is a widely used measure of depression and has good, well-established psychometric properties. The scales indicate good internal consistency (Cronbach coefficient alpha = .93 for college students) and correlate well with several clinical assessment ratings of depression (Beck et al., 1996).

The *STAI Form Y-2* is used to assess levels of trait anxiety in both clinical and

non-clinical populations. It comprises of 20 self-statements that ask individuals to rate how they feel in general, on a scale of 1 (almost never) to 4 (almost always). The scales contain nine negatively reversed items. Total scores range from 20 (indicating low levels of general anxiety) to 80 (indicating high levels of general anxiety -- mean score for college-aged males = 38.3,  $SD = 9.2$  and females:  $M = 40.40$ ,  $SD = 10.15$ ). The STAI Form Y-2 is a valid measure of anxiety indicating good test-retest reliability for a sample of college students (.73 to .86) and high internal consistency (Cronbach coefficient alpha is .90). The STAI Form Y-2 is also reliable and correlates with other established anxiety measures (see Spielberger et al., 1983).

*Self-statements.* The automated procedure employed in Experiment 1 involved the presentation of a series of single-sentence self-statements, half of which were referred to as *positive* (e.g. "I am whole") and half referred to as *negative* (e.g. "I am stupid"). In order to generate two fitting sets of statements that would be readily categorised by participants in this way, a number of ACT therapists were asked to generate a list of statements that resemble the types of self-content with which clients typically struggle. The resulting list comprised of 40 statements (20 positive and 20 negative), which was subsequently presented to 32 independent raters (none recruited as experimental participants). These were asked to rate how positive or negative each statement was on a Likert scale of 1-21, where 1 was extremely negative and 21 was extremely positive. The mean ratings for all 40 statements are presented in Appendix D. The 10 self-statements rated as the *most positive* and the 10 rated as the *most negative* were selected for presentation in the experiment. The mean ratings for the 10 positive statements ranged from 18.66 ("I love life") to 16.37 ("I am whole"). The mean ratings for the 10 negative statements ranged from 1.45 for "I am a failure" to 3.23 for "No-one will ever love me".

*Automated Procedure.* The automated procedure was always conducted on a PC laptop with a Pentium III processor and a 12" LCD screen. All stimulus presentations and participant responses were recorded by a Visual Basic (VB) program (Version 6). A standard computer mouse was also used to enable participants to interact directly with the program.

*Adherence Measures.* At the end of the automated procedure, participants were asked to complete two 7-point Likert scales in paper-and-pen format. These were employed as adherence measures and asked participants: (1) "To what extent did you read the statements and ratings and answer them honestly?" (referred to as the *Honesty Scale*) and (2) "To what extent, do you think your ratings have been affected by the instructions you received at the beginning of the study?" (referred to as *Control by defusion-related instructions*). Each rating scale ranged from 1 (not at all) to 7 (entirely).

### *Experimental Overview*

This study used a 3x3x2 mixed-between-within participants design, with instructions (regarding the impact of defusion as a coping strategy) as the between participant variable (Pro-Defusion, Anti-Defusion and Neutral), and visual presentation format of the self-statements (Normal, Defused and Abnormal) plus statement type (positive and negative) as within-participant variables.

In Phase 1, participants simply completed the AAQ-35, the BDI-II and the STAI-Y2. In Phase 2, they were randomly assigned to one of three groups that were differentiated in terms of the instructions they received regarding defusion (Pro-Defusion, Anti-Defusion and Neutral). Specifically, participants in the *Pro-Defusion Group* received instructions that indicated that defusion *decreases* the emotional



impact of self-statements. Participants in the *Anti-Defusion Group* were instructed that defusion *increases* the emotional impact of self-statements. And participants in the *Neutral Group* were informed that defusion has *no effect* on the emotional impact of self-statements.

In Phase 3, all participants were exposed to the same automated procedure that involved the presentation of the positive and negative self-statements. Each statement was presented in three different presentation formats (Normal, Defused and Abnormal) to determine the potential impact of the defusion prefix relative to the other presentations. In the *Normal* presentation format, each self-statement appeared without a prefix (e.g. “I am a bad person”). In *Defused* presentation, each statement was prefixed with “I am having the thought that” (e.g. “I am having the thought that I am a bad person”). In *Abnormal* presentation, each statement was prefixed with “I have a wooden chair and” (e.g. “I have a wooden chair and I am a bad person”). Immediately after each statement, participants rated their levels of discomfort, believability and willingness regarding the statement. Phase 4 simply comprised the adherence measures.

### *Ethical Issues*

The current study raised a number of ethical considerations, in light of the fact that potentially negative psychological consequences could result, particularly from exposure to the negative self-statements. In order to conduct the study within all of the appropriate ethical guidelines (The Psychological Society of Ireland, 2003), a series of precautionary measures was taken with all participants. First, each participant was briefed as to the nature of the study *prior to* agreement to participate. If, at this point, participants agreed to proceed, they were requested to sign a consent

form (Appendix E) that was then counter-signed by the experimenter. All participants were informed at this point that they were not required to proceed; that they could withdraw from the study *at any point*; and that they would not be contacted again.

Second, prior to the three psychological measures, participants were informed that the three questionnaires contained questions that they might find personal and/or distressing. In order to address potential anxieties concerning questionnaires, it was emphasised that: (a) participants could avoid answering any question; (b) there were no right or wrong answers to the questions; and (c) that the questionnaires served merely to enable the experimenter to note individual differences that might influence performances during the experimental procedure. At this point, participants were also informed that they would be provided with their own overall scores (without interpretation) on the questionnaires, but not the results of other participants.

Third, participants were informed that as part of the experimental procedure they would be exposed to a number of positive and negative self-statements, some of which they may find distressing. Fourth, participants were informed that, in accordance with guidelines, all data would be kept on file at the Psychology Department in NUIM for approximately 5 years. Furthermore, they were assured that all aspects of their participation in the study would remain completely confidential (i.e. their data would not be identified by name in the final report and all data would be coded with an individual identification code number). Data would only be disclosed to the thesis supervisor in the event of exceptional circumstances (e.g. in the event of a participant becoming distressed or expressing a wish to see the supervisor). Fifth, after their participation in the study, all participants were debriefed on the nature and purpose of the study. Specifically, the concept of cognitive defusion was explained to them and the reasons for the use of the Abnormal self-statements as well

as the different conditions were outlined. Sixth, all participants were made aware that if any psychological issues arose *directly* from the study, they could contact the thesis supervisor (whose name was provided) and further arrangements, where appropriate, could be made free of charge as a forum for the discussion of these issues with a Chartered Psychologist. At no point during the experiment did any participant withdraw from the study or express dissatisfaction or distress of any kind.

### *Procedure*

*Phase 1: Psychological Measures.* During Phase 1, participants completed the AAQ-35, the BDI-II and the STAI-Y2, presented in written booklet format in this order. The assignment of participants to experimental conditions was randomised and did not depend upon their scores on the questionnaires. The three measures were simply included as a means of determining any differences between the three experimental groups that might influence their performance during the experiment. For example, participants who scored high in acceptance (i.e. low in avoidance) on the AAQ-35 might respond more readily to defusion than those who scored low in acceptance (high in avoidance). Furthermore, the latter may have been less willing to engage with the negative self-statements and this might have considerably undermined their full participation in the study.

*Phase 2: Presentation of Therapeutic Rationales.* Each participant was randomly assigned to one of the three experimental groups (Pro-Defusion, Anti-Defusion and Neutral) prior to completing Phase 1. Each group was presented with similar instructions, but with minor differences for each group (although the overall length of the instruction sets was almost identical).

Participants in the *Pro-Defusion Group* were provided with a set of instructions designed to encourage them *towards* a defusion rationale. The instructions were printed on a sheet of A4 paper and remained available to participants for the duration of the experiment. These were as follows:

In the current experiment, we are interested in the emotional impact of unusual self-statements. The scientific literature in this area shows that if you rephrase a self-statement like “I am an awful person” into “I am having the thought that I am an awful person”, then the emotional impact of the statement is *reduced*.

In other words, thinking or saying words like “I am having the thought that I am an awful person” is *easier* to deal with than simply thinking or saying “I am an awful person”.

Immediately after reading the instructions, participants were provided with a brief adherence measure in written format, designed to determine whether or not they had fully understood the instructions. During this measure, participants were presented with the following instructions:

Just to check that you understand the previous paragraph, please answer the following question and tick ONE of the choices below.

The statement “I am having the thought that I am an awful person” should INCREASE/DECREASE/NOT AFFECT your emotional reaction to the statement “I am an awful person”.

Participants were simply required to tick the choice they felt to be most appropriate to the instructions they had received. In the case of the Pro-Defusion Group, therefore, the correct answer involved indicating that the prefix “I am having the thought that” would *decrease* one’s emotional reaction to the statement as normally presented.

Participants in the *Anti-Defusion Group* were provided with an almost identical set of instructions, except that they were encouraged *away from* (rather than *towards*) a defusion strategy. These instructions were as follows:

In the current experiment, we are interested in the emotional impact of unusual self-statements. The scientific literature in this area shows that if you rephrase a self-statement like “I am an awful person” into “I am having

the thought that I am an awful person”, then the emotional impact of the self-statement is *increased*.

In other words, thinking or saying words like “I am having the thought that I am an awful person” is *harder* to deal with than simply thinking or saying “I am an awful person”.

Immediately after reading the instructions, participants in the Anti-Defusion group were also required to complete a brief adherence measure to determine whether or not they had fully understood the instructions. In the case of the Anti-Defusion Group, the correct answer involved indicating that the prefix “I am having the thought that” would *increase* one’s emotional reaction to the statement normally presented.

The instructions provided to the *Neutral Group* informed participants that a defusion prefix would have *no effect* on one’s emotional reaction to the self-statement. That is, they were neither encouraged towards, nor away from, a defusion strategy. In this way, the Neutral group was employed as a type of experimental control condition and participants were instructed as follows:

In the current experiment, we are interested in the emotional impact of unusual self-statements. The scientific literature in this area shows that if you rephrase a self-statement like “I am an awful person” into “I am having the thought that I am an awful person”, then the emotional impact of the statement *does not change*.

In other words, thinking or saying words like “I am having the thought that I am an awful person” is *no different than* simply thinking or saying “I am an awful person”.

Participants in the Neutral group were also required to complete a brief adherence measure in which the correct answer involved indicating that the prefix “I am having the thought that” would *not effect* one’s emotional reaction to the statement normally presented. Having successfully completed Phase 2 of the experiment, all participants were invited to ask any questions regarding any aspect of the instructions they had received thus far.

*Phase 3: Automated Presentation and Ratings of Self-Statements.* Irrespective of the experimental group to which participants had been assigned, all were exposed to an identical automated procedure that contained 60 individual presentations of self-statements (i.e. three different exposures to each of the 10 positive statements and three different exposures to each of the 10 negative statements). The three types of exposure to each of the self-statements were designed in an attempt to manipulate explicitly the level of defusion created for each participant by each statement *as it appeared on screen*. (This second defusion manipulation was in addition to the manipulation of defusion via the experimental instructions presented in Phase 2). Each of the three exposures to each statement was referred to as *Normal*, *Defused* and *Abnormal*, and simply reflected the format in which the statement was presented. All 60 automated trials were randomised by the program for each participant. Exposure to the automated procedure lasted approximately 40-50mins.

During the *Normal presentation format*, each self-statement appeared in its original form. That is, for example, the self-statement “I am a bad person” appeared on screen in exactly this form. Statements presented in the *Defused presentation format* were formed by placing the prefix “*I am having the thought that*” in front of each statement. Thus, for example, the self-statement “I am a bad person” appeared on screen as “I am having the thought that I am a bad person” when presented in the Defused format. During the *Abnormal presentation format*, each self-statement was modified by adding the unusual prefix “I have a wooden chair and” in front of the statement (e.g. “I have a wooden chair and I am a bad person”). This modification acted as a type of experimental control for the difference in length between the Normal and Defused self-statements. That is, the Abnormal statements were the same length as the Defused statements, but contained no defusion content. Potential

differences, therefore, in ratings for the Normal and Defused self-statements that resulted from word length should yield no differences between the Abnormal and Defused statements.

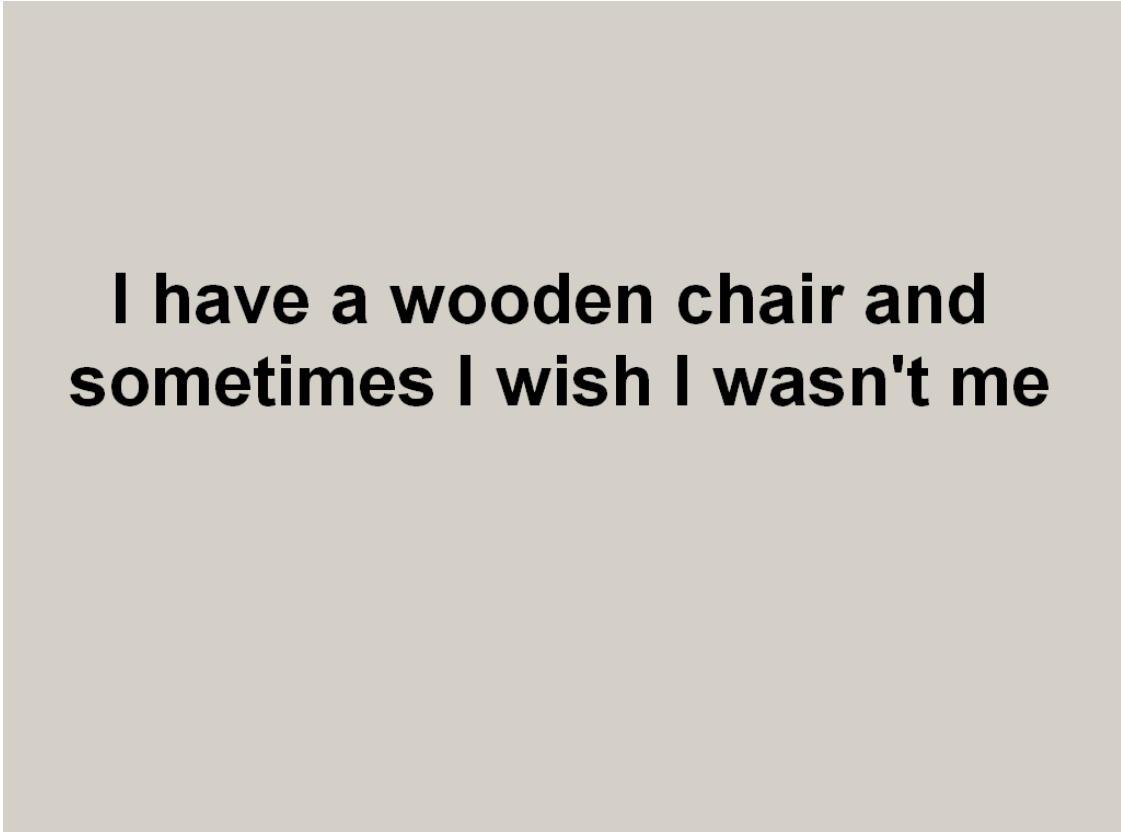
Prior to the commencement of the experimental trials, participants were exposed to two automated practice trials that were identical in format to the experimental trials, but did not contain any of the target statements. The program commenced with an introductory screen that consisted of a blue background and a grey button box in the centre that said: “*Click Here When You Are Ready to Start the Practice Trials*”. All interactions with the program involved using the mouse to move and click the cursor on-screen. When participants were ready to continue, they simply clicked the button box and a grey instruction screen appeared automatically. At the top of this was the instruction: “*Read the following statement carefully and think about it*”. This remained on-screen for 3s., after which a self-statement appeared automatically in the centre of the screen for 6s. This was followed by a rating screen that contained three individual rating scales (that represented discomfort, believability and willingness – see below), each ranging from 0-100, presented against a grey background. When participants had completed all three rating scales, they were instructed to click on the green button box at the bottom centre of the screen. When they placed the cursor on the box and clicked, the instruction screen appeared again and a second statement and rating screen followed. When the second practice ratings were provided, the automated procedure began in earnest. This marked the end of the practice session. Immediately after this, the automated procedure commenced with the introductory screen.

The automated procedure contained a series of instructions throughout that allowed participants to proceed correctly through all aspects of the program in the

absence of the experimenter. All of these instructions appeared in black letters and were clearly visible to participants. The automated procedure consisted of a total of 242 screens: an introductory screen; 60 instruction screens that preceded each of the 60 self-statements; 60 intermediate screens that followed each of the 60 rating screens; and an end screen that marked the end of the program.

The introductory screen that preceded the experimental trials consisted of a blue background, with a grey button box in the centre that said: “*Click Here When You Are Ready to Start the Experiment*”. When participants were ready to continue, they simply located the cursor on the button box and clicked. The first instruction screen then appeared automatically. This consisted of a grey background, at the top of which was written the instruction: “*Read the following statement carefully and think about it*”. This instruction remained on-screen for 3s. Each of the 60 self-statement screens was preceded by an identical instruction screen. After 3s., one of the self-statements appeared automatically in the centre of the screen against a grey background. The statement remained there for 6s. An example of a negative self-statement in the Abnormal presentation format is illustrated in Figure 1. Participants were not required to make any response in the presence of the self-statement.





**I have a wooden chair and  
sometimes I wish I wasn't me**

*Figure 1.* An example of a negative self-statement in Abnormal presentation format in Experiment 1.

Each self-statement screen was followed by a rating screen that contained three individual rating scales, ranging from 0-100, and presented against a grey background (see Figure 2). At the top centre of the screen was the generic instruction: *“Please Use the Sliding Scales to Answer Each of the Following Questions”*. Immediately below, there were three indented rating boxes in which participants were required to provide ratings of discomfort, believability and willingness for the previous self-statement.

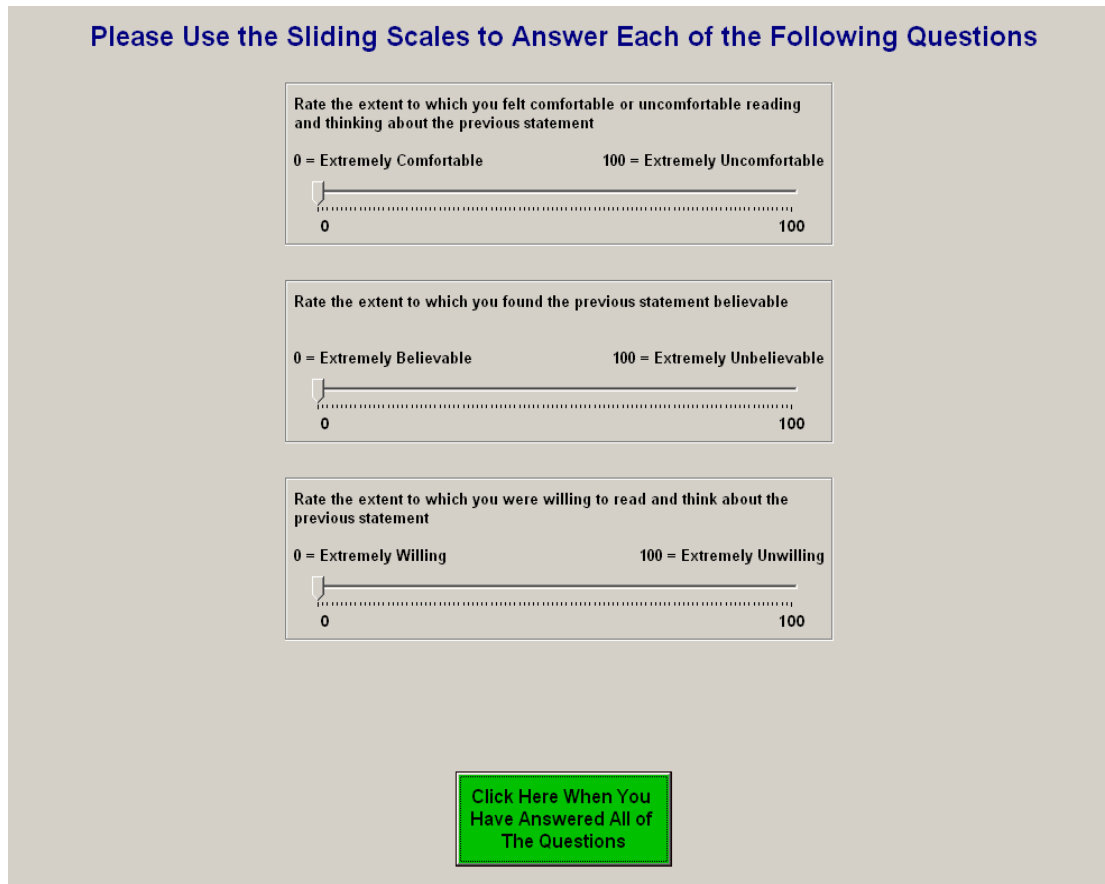


Figure 2. The experimental rating screen that followed the presentation of each automated self-statement in Experiment 1.

The Discomfort rating scale appeared at the top with the following instruction: *“Rate the extent to which you felt comfortable or uncomfortable reading and thinking about the previous statement”*. The sliding scale below was marked “0 = Extremely Comfortable” on the left and “100 = Extremely Uncomfortable” on the right. To provide a rating, participants could place the cursor on the slider, hold the left mouse button down, and drag the cursor to the appropriate point on the scale. Alternatively, they could click directly on the rating point with the cursor, until the slider moved. The initial slider value for each scale was set at -1 in order to force participants to move the cursor for each of the three scales before proceeding to the next screen.

The Believability scale appeared immediately below and contained the

instruction: *“Rate the extent to which you found the previous statement believable”*. In this case, the extreme ends were marked with “0 = Extremely Believable” on the left and “100 = Extremely Unbelievable” on the right. The Willingness scale was at the bottom and the instruction was: *“Rate the extent to which you were willing to read and think about the previous statement”*. In this case, ratings ranged from “0 = Extremely Willing” to “100 = Extremely Unwilling”.

At the bottom of each rating screen was a green button box with the following instructions: *“Click Here When You Have Answered All of The Questions”*. When participants clicked on the box, an intermediate screen appeared immediately. This screen consisted of a grey background, with a grey button box in the centre, on which it was written *“Click Here to Continue”*. These screens allowed participants to proceed through the experimental trials at their own pace. When participants clicked the grey box, another self-statement screen appeared immediately. This sequence of screens continued until all 60 statements had been presented and rated.

When participants had completed the final rating screen, an ‘end screen’ appeared immediately. This consisted of a blue background, with three pale blue dialogue boxes located one above the other in the centre of the screen. The box at the top contained the following: *“This is the end of this part of the experiment”*. The box in the middle of the screen contained the words: *“Thank you for your participation”*. Finally, the box at the bottom of the screen contained the words: *“Please Report to the Experimenter”*. The end screen marked the end of the automated procedure. At this point, participants were required to seek the attention of the experimenter who was seated outside of the room.

*Phase 4: Adherence Measures.* Immediately after the presentation of the 60 self-statements, participants were required to complete the two adherence questions.

These were presented on a piece of A4 paper and comprised of two rating scales, referred to for current purposes as (1) the Honesty Scale and (2) Control by defusion-related instructions. At the end of Phase 4, participants were thanked and debriefed. Each aspect of the study was conducted in a single session and each individual took approximately 90mins. to complete.

## RESULTS

### *Psychological Measures*

The mean ratings (and standard deviations) for each of the psychological measures (AAQ-35, BDI-II and STAI-Y2) for each group of participants are presented in Table 1. The results from the BDI-II and the STAI-Y2 indicated that participants fell within the normal range for depression and anxiety respectively. Results from the AAQ-35 indicated that participants overall reported high levels of acceptance/low levels of avoidance. In each case, the measures were broadly similar across the three instructional groups. Indeed, three separate one-way between groups analyses of variance (ANOVA's), one for each of the psychological measures, revealed no significant differences across groups (AAQ-35: [ $p = .67$ ]; BDI-II: [ $p = .25$ ]; STAI-Y2: [ $p = .84$ ]). Subsequent differences, therefore, in response to any aspect of the experiment could not readily be attributed to pre-experimental variations across the groups on any of these measures.

Table 1

*Means (and Standard Deviations) for Psychological Measures for Each Instruction Group in Experiment 1*

	Norms	Pro-Defusion	Anti-Defusion	Neutral
AAQ-35	-	160.95 (17.37)	158.25 (12.54)	163.25 (21.81)
BDI-II	< 13	7.25 (4.87)	6.10 (5.77)	8.90 (5.26)
STAI-Y2	40.40 (10.15)	36.45 (8.42)	37.25 (11.15)	38.40 (11.31)

*Note.* ‘-’ indicates that no normative data is available.

*Adherence Measures*

High ratings were obtained on the Honesty Scale (maximum score = 7) across all three groups (Pro-Defusion:  $M = 6.20$ ,  $SD = .61$ ; Anti-Defusion:  $M = 5.95$ ,  $SD = .99$ ; and Neutral:  $M = 5.90$ ,  $SD = .85$ ), indicating that participants on the whole responded honestly to the self-statements. A one-way between groups ANOVA revealed that there were no significant differences between groups [ $F(2, 57) = .74$ ,  $p = .48$ ] in this regard.

In contrast, the mean ratings on the Control by defusion-related instructions Scale (maximum = 7) appeared to be relatively low for each of the three groups (Pro-Defusion:  $M = 3.40$ ,  $SD = 1.90$ ; Anti-Defusion:  $M = 2.95$ ,  $SD = 1.98$ ; and Neutral:  $M = 3.90$ ,  $SD = 2.07$ ), indicating that ratings were not affected by the defusion-related instructions at the beginning of the study. Perhaps more importantly however, a one-way between groups ANOVA revealed no significant difference between the groups [ $F(2, 57) = 1.14$ ,  $p = .33$ ]. Thus, although participants were not greatly influenced by the defusion-related instructions, all three groups were equal in this regard.

### *Self-Report Measures*

For the purposes of analysis, the positive and negative self-statements were divided according to the three presentation formats: Normal, Abnormal and Defused. The three ratings (discomfort, believability and willingness) for each of the three formats, summed across the ten trials for each format, ranged from 0 to 1000 (e.g. for discomfort ratings of the negative statements in normal format; 10 statements, each of which were rated on a scale of 0 to 100, providing a possible total score of 0 to 1000). The data obtained for the positive and negative self-statements were analysed separately and are presented in separate sections below.

### *Negative Self-Statements*

The mean overall scores for the negative self-statements were calculated for each of the three rating scales (Discomfort, Believability and Willingness), for each of the three groups (Pro-Defusion, Anti-Defusion and Neutral). These scores were then separated into the three presentation formats (Normal, Abnormal and Defused -- see Table 2).

Table 2

*Means (and Standard Deviations) for Discomfort, Believability and Willingness of the Three Presentation Formats for the Three Groups for Negative Self-Statements for Experiment 1*

Presentation Format	Pro-Defusion Group	Anti-Defusion Group	Neutral Group
<i>Discomfort ratings</i>			
Normal	326.80 (293.18)	465.65 (322.41)	390.95 (268.59)
Abnormal	284.70 (261.68)	421.50 (275.46)	412.75 (293.69)
Defused	219.90 (220.62)	393.70 (297.62)	323.05 (232.14)
<i>Believability ratings</i>			
Normal	849.25 (99.66)	709.50 (264.92)	837.10 (132.21)
Abnormal	835.05 (145.51)	706.35 (259.01)	817.25 (149.89)
Defused	630.80 (199.70)	625.90 (270.13)	726.00 (127.79)
<i>Willingness ratings</i>			
Normal	347.85 (336.91)	453.40 (314.58)	481.80 (266.26)
Abnormal	346.95 (325.22)	433.65 (255.67)	505.75 (286.82)
Defused	245.30 (259.04)	369.65 (281.98)	400.25 (262.42)

*Note.* Lower scores indicate lower discomfort, greater believability and greater willingness.

The descriptive statistics overall presented positive outcomes for both the Pro-Defusion instructions and the Defused presentation format. Specifically, within each instruction group, the Defused presentation format produced lowest levels of discomfort and highest levels of willingness for the negative self-statements, relative to Normal and Abnormal presentations. Furthermore, across group comparisons indicated that the Pro-Defusion instructions were also associated with the lowest levels of discomfort and interestingly the Anti-Defusion instructions were associated with greatest discomfort. Similarly, the Pro-Defusion instructions produced greatest willingness (with Neutral instructions generating lowest willingness). Contrary to predictions however, the Defused format was associated with highest levels of believability, although the Pro-Defusion instructions did not yield the highest believability ratings compared to the other groups. The statistical analyses conducted for each of the three types of ratings are discussed separately below.

*Discomfort Ratings.* A 3x3 mixed-between-within repeated measures ANOVA was conducted on the discomfort ratings with instruction group (Pro-Defusion, Anti-Defusion and Neutral) as the between participant variable and presentation format (Normal, Abnormal and Defused) as the within participant variable. A significant main effect was identified for presentation format ( $F(2, 57) = 19.05, p < .0001, \eta_p^2 = .25$ ), but not for group ( $p = .21$ ) and the interaction was also non-significant ( $p = .17$ ). Post-hoc (Scheffe) tests revealed a significant difference between Defused and Normal presentations ( $p < .0001$ ), and between Defused and Abnormal ( $p = .0001$ ), but not between Normal and Abnormal presentations. In short, defusion significantly decreased discomfort relative to the other two presentation formats and this effect was not significantly modulated by the defusion-related instructions. The fact that the Defused format differed significantly from both Normal and Abnormal, but the latter did not differ from one another suggested that the effect of the Defused format on discomfort could not be accounted for in terms of the length or unusual nature of the defused statements.

*Believability Ratings.* The 3x3 ANOVA on the believability data also revealed a significant main effect for presentation format ( $F(2, 57) = 44.30, p < .0001, \eta_p^2 = .43$ ), but not for group ( $p = .12$ ). In this case however, the interaction was significant ( $F(4, 114) = 4.18, p = .003, \eta_p^2 = .12$ ). In order to examine this effect, three separate one-way repeated measures ANOVAs were conducted, one for each group. The Pro-Defusion instruction group showed a significant effect for presentation format ( $F(2, 19) = 24.19, p < .0001, \eta_p^2 = .56$ ) and Sheffe post-hoc tests indicated that the Defused format differed significantly from both Normal and Abnormal presentations ( $p < .0001$  and  $p < .0001$ , respectively), with no significant difference between the latter ( $p = .92$ ). The same pattern was recorded with the Anti-Defusion group, who also



showed a significant effect for format ( $F(2, 19) = 6.33, p = .004, \eta_p^2 = .25$ ), with Defused again significantly different from both Normal and Abnormal ( $p = .0125$  and  $p = .0168$ , respectively), and no significant difference between the latter ( $p = .99$ ). The case was the same for the Neutral group, for whom format was also significant ( $F(2, 19) = 16.98, p < .0001, \eta_p^2 = .47$ ), and Defused was significantly different from both Normal and Abnormal ( $p < .0001, p = .0003$ , respectively), but with no significant difference between the latter ( $p = .62$ ). Contrary to predictions therefore, the Defused presentation format significantly *increased* believability relative to the Normal and Abnormal presentations for each of the three groups.

*Willingness Ratings.* Another 3x3 ANOVA was conducted on the willingness ratings. Similar to the discomfort measure, a significant main effect was recorded for presentation format ( $F(2, 57) = 20.91, p < .0001, \eta_p^2 = .26$ ), but not for group ( $p = .23$ ), and again the interaction was non-significant ( $p = .74$ ). Scheffe post-hoc tests revealed the same pattern as discomfort (Defused vs. Normal =  $p < .0001$ ; Defused vs. Abnormal =  $p < .0001$ ; Normal vs. Abnormal *ns*). In effect, defusion significantly increased willingness relative to the other two presentation formats, with no modulation by the defusion-related instructions, and no difference between the Normal and Abnormal statements. The fact that the Defused format differed significantly from both Normal and Abnormal, but the latter did not differ from one another suggested that the effect of the Defused format on willingness could not be accounted for in terms of the length or unusual nature of the defused statements.

### *Positive Self-Statements*

The mean overall scores for the positive self-statements were calculated for each of the three rating scales (Discomfort, Believability and Willingness), for each of the three groups (Pro-Defusion, Anti-Defusion and Neutral). These scores were then separated into the three presentation formats (Normal, Abnormal and Defused -- see Table 3).

Table 3

*Means (and Standard Deviations) for Discomfort, Believability and Willingness Ratings of the Three Presentation Formats for the Three Groups for Positive Self-Statements in Experiment 1*

Presentation Format	Pro-Defusion Group	Anti-Defusion Group	Neutral Group
<i>Discomfort ratings</i>			
Normal	112.55 (101.09)	141.40 (126.61)	158.25 (153.05)
Abnormal	122.15 (117.04)	189.90 (155.74)	220.55 (173.78)
Defused	119.40 (137.72)	195.50 (147.08)	166.95 (142.43)
<i>Believability ratings</i>			
Normal	196.80 (140.65)	196.30 (143.53)	227.75 (156.34)
Abnormal	249.35 (191.89)	301.75 (190.77)	278.60 (125.09)
Defused	218.95 (204.13)	246.50 (149.09)	258.40 (156.77)
<i>Willingness ratings</i>			
Normal	125.05 (118.75)	172.25 (144.99)	216.95 (175.51)
Abnormal	162.15 (178.62)	227.95 (184.08)	285.10 (198.43)
Defused	135.15 (157.85)	190.85 (154.00)	219.25 (158.75)

*Note.* Lower scores indicate lower discomfort, greater believability and greater willingness.

In general, the pattern of ratings associated with the positive self-statements differed considerably from those recorded with the negative statements. Specifically within each group, the Normal presentation format was associated with greatest willingness, greatest believability and lowest discomfort. In contrast, the across group comparisons indicated that the Pro-Defusion instructions were associated with lowest discomfort and greatest willingness. The pattern of responses on believability was

generally unsystematic across groups, although there was some indication that Pro-Defusion was associated with greatest believability.

*Discomfort Ratings.* A 3x3 mixed-between-within repeated measures ANOVA, with instruction group as the between participant variable and presentation format as the within participant variable, revealed a significant main effect for format [ $F(2, 57) = 6.64, p = .002, \eta_p^2 = .10$ ], but not for group ( $p = .25$ ). The interaction was non-significant ( $p = .07$ ). Post-hoc (Scheffe) tests indicated a significant difference only between the Normal and Abnormal statements ( $p = .002$ ), while Defused did not differ significantly from either -- Normal ( $p = .12$ ) and Abnormal ( $p = .31$ ). Unlike the negative statements therefore, the Defused presentation format had little impact on the discomfort ratings obtained with the positive statements. Once again, the defusion-related instructions also had little or no effect on levels of discomfort.

*Believability Ratings.* The 3x3 ANOVA conducted on the positive believability ratings also revealed a significant main effect for presentation format [ $F(2, 57) = 9.79, p = .0001, \eta_p^2 = .15$ ], but not for group ( $p = .76$ ), and the interaction was non-significant ( $p = .62$ ). Similar to the discomfort ratings, post-hoc (Scheffe) tests revealed a significant difference between Normal and Abnormal formats only ( $p = .0001$ ), but not between Defused and Normal ( $p = .10$ ), or between Defused and Abnormal ( $p = .09$ ). In short, the Defused presentation format and the defusion-related instructions had little impact on the believability ratings of the positive statements.

*Willingness Ratings.* The 3x3 ANOVA for willingness also revealed a significant effect for presentation format [ $F(2, 57) = 9.54, p = .0001, \eta_p^2 = .14$ ], but not for group ( $p = .13$ ). Again, the interaction effect was non-significant ( $p = .76$ ).

Interestingly, post-hoc (Scheffe) tests revealed a significant difference between the Normal and Abnormal statements ( $p = .0004$ ) and between Defused and Abnormal ( $p = .005$ ), but Normal and Defused did not differ from one another ( $p = .73$ ). Thus, participants were more willing to experience the positive statements as Normal or Defused than as Abnormal. Once again, the defusion-related instructions had little impact.

In summary, a different pattern of responding was observed for the negative versus positive self-statements. For the negative statements, the Defused presentation format produced significantly lowest levels of discomfort, highest willingness and highest believability relative the Normal and Abnormal presentations. In contrast, this type of presentation had relatively little impact on the emotional ratings associated with the positive statements. Overall, the three different types of defusion-related instructions had little or no impact on the emotional ratings obtained with either negative or positive statements presented in any of the three presentation formats.

## DISCUSSION

The results obtained for Experiment 1 were consistent with those reported by Masuda et al. (2004) in terms of the types of differences in the emotional impact of negative self-statements reported by participants when presented with statements involving defusion, compared to those not-defused. In both studies, defusion involving negative self-statements was associated with decreases in discomfort. Across the two studies however, a key difference emerged in the believability ratings. In the original work, defusion was associated with decreased believability, and although the same prediction was made here, the current research actually indicated

that defusion was associated with an *increase* in believability.

The current study offered a number of additional advantages over the original work that potentially shed light on explanations of both data sets. (1) The current study investigated the impact of defusion on willingness to experience the negative statements, and as expected defusion was associated with increased willingness. (2) This work was the first to examine the impact of defusion using the defusion prefix (“I am having the thought that”), rather than a word repetition technique. (3) This work systematically compared the impact of defusion on positive versus negative self-statements and indicated that defusion had relatively little impact on the emotional ratings associated with the former. (4) Finally, the current study offered a more systematic manipulation of defusion in terms of the relative impact of defusion instructions versus a defusion exercise. Consistent with the original research, the current work also found that defusion-related instructions were considerably less influential than a more direct defusion manipulation.

Overall, the research presented here provides empirical evidence of the utility of cognitive defusion in reducing the emotional impact of negative self-statements, while encouraging willingness to experience them. Although there were points of overlap with the only existing published study in this area (e.g. discomfort ratings and the utility of a defusion experience versus instructions), there were also points of inconsistency (e.g. the believability ratings). Taken together, this pattern raised a number of important questions for future research in this area.

First, participants here were exposed to generic self-statements, whereas personalised statements had been employed previously. Thus, it was difficult to determine what impact this difference may have had on the two sets of findings. Second, the defusion-related instructions and defusion exercises were brief, and one

might even argue that the manipulation of the presentation formats of the statements employed here cannot adequately be called ‘an intervention’. Hence, it might be argued that different outcomes would be obtained with a more substantive defusion-based exercise. Such an intervention would also be more in line with the types of defusion techniques commonly employed in therapeutic settings. Third, although the original research by Masuda et al. (2004) examined the comparison between defusion and distraction-based and thought control strategies, no such comparison was attempted here. The next phase of the current research program was designed to address these issues directly. Specifically, Experiment 2 attempted to: (1) investigate the impact of defusion on personally relevant thoughts; (2) present participants with a more complete therapeutic defusion-based rationale and exercise; and (3) compare the impact of defusion with alternative coping strategies. These manipulations allowed for a more direct comparison with the Masuda et al. study.

## **Chapter 3**

### **Experiment 2**

**Cognitive Defusion and Personally Relevant**

**Negative Thoughts:**

**Exploring the Relative Impact of Therapeutic**

**Instructions and Experiential Exercise**

## **Chapter 3**

### **Experiment 2**

# **Cognitive Defusion and Personally Relevant Negative Thoughts: Exploring the Relative Impact of Therapeutic Instructions and Experiential Exercise**

## INTRODUCTION

Experiment 1 investigated the emotional impact of defusion instructions and a defusion exercise on positive and negative self-statements. The results obtained were largely consistent with existing empirical evidence in spite of a number of methodological differences between this and the only previously published study (Masuda et al., 2004). First, both studies had attempted to compare the relative impact of defusion when delivered via instruction versus experiential exercise and the results from both generally indicated superiority of exercise over instruction in this regard. However, while the defusion instructions presented by Masuda et al. to participants were relatively lengthy and detailed, those provided in Experiment 1 were extremely brief. Furthermore, the latter could not easily be called a clinical analogue because it is unlikely that therapists working within an ACT framework would provide such brief instructions to clients in introducing defusion, even when they are to be followed by experiential exercises.



In the original research, Masuda et al. described a condition in which participants were presented with defusion instructions and were then simply instructed to apply this strategy to their negative thoughts, with no opportunity to practice defusion with a thought. Although this appears to offer a condition in which instructions were presented in the absence of experiential exercise, one cannot determine what participants actually did with the information they received. For example, some participants may have successfully understood the implications of the rationale and applied defusion appropriately to their negative thought, others may not. Furthermore, at no point did the researchers investigate the impact of the defusion exercise alone. In Experiment 1 of the current research, all participants received both defusion-related instructions and exercise. As a result, it might be argued that neither study addressed the relative impact of defusion instructions versus exercise.

A related issue emerges with the thought control strategy employed by Masuda et al. (2004). The researchers provided participants with a *number of different* strategies that they might use (e.g. positive self-talk, positive imagery, or breathing training), hence it is difficult to determine which strategy was actually used and which is most effective. Second, a key difference between the studies concerned the use of personally relevant self-statements employed by Masuda, but not in Experiment 1. Although both studies yielded similarly positive outcomes for defusion, it remains difficult to determine the role played by the use of personal statements. In any case, it is reasonable to argue that using personal statements offers a better clinical analogue of what happens in natural therapeutic settings, thus this is an important issue to examine experimentally.

Experiment 2 attempted to address these issues in a number of ways. First, we investigated the impact of defusion on participants' personally relevant negative thoughts. Second, we systematically examined the relative utility of defusion instructions versus defusion exercises. Third, we compared defusion with other coping strategies, including Thought Control and Placebo. Fourth, we specifically targeted one type of Thought Control strategy which all participants were trained to use. In this way, Experiment 2 offered a more direct comparison with the Masuda et al. (2004) study.

Experiment 2 also incorporated a number of additional features designed to extend both the original research by Masuda et al. (2004) and Experiment 1. Specifically, rather than using a single participant alternating treatment design employed by Masuda, the current study adopted an independent groups design to permit greater generalisation of the findings. To control for possibly confounding effects of demand compliance (e.g. Field, & Lawson, 2003), all experimental aspects of the research were conducted in the absence of the experimenter, with all instructions presented in written format. Finally, all 'interventions' employed in Experiment 2 were considerably more substantive than those employed previously.

### *The Current Study*

In Experiment 2, a sample of non-clinical undergraduates was randomly assigned across nine experimental conditions. Each condition consisted of the presentation of one of three therapeutic rationales (Defusion, Thought Control, or Placebo) paired with one of three experiential exercises (Defusion, Thought Control, or Placebo). Thus, a condition may have consisted of *matched* components, where both components comprised the same

therapeutic strategy (e.g. a defusion rationale plus a defusion exercise), or both contained no strategy (a placebo rationale combined with a placebo exercise). A condition may also have consisted of *mismatched* components, where one component offered one strategy, while the other component offered the alternative strategy (e.g. a defusion rationale paired with a thought control exercise). Finally, some conditions were also referred to as *unmatched*, where one component involving a strategy was accompanied by a placebo component (e.g. a defusion rationale paired with a placebo exercise). Participants were required to generate a negative self-relevant thought (e.g. “I am a failure”) and rate this thought in terms of the discomfort, believability and willingness it induced prior to, and immediately after, the presentation of the therapeutic rationale and exercise.

Consistent with both Experiment 1 and Masuda et al. (2004), we hypothesised that participants presented with defusion components would demonstrate decreases in the discomfort. Consistent with the work presented by Masuda et al., we also hypothesised that defusion would decrease believability. Finally, in line with the research presented in Experiment 1, we hypothesised that defusion would increase willingness (Masuda et al. did not investigate willingness). We also anticipated that participants presented with two defusion components (i.e. both the Defusion rationale and Defusion exercise) would produce even greater improvements. Furthermore, we expected that the Defusion exercise would likely prove to be of greater impact than the Defusion rationale.

## METHOD

### *Participants*

One hundred and forty-four undergraduates (72 male) aged between 18 and 22 years ( $M = 20.6$  years) were recruited for Experiment 2 through Faculty announcements at NUIM. None had prior exposure to similar experimental procedures. Participants were randomly assigned to one of nine conditions, each of which was balanced for gender (8 males and 8 females per condition). No incentives were offered in return for participation. In addition to the 144 participants who completed the experiment, three failed to complete all three psychological measures and a further two failed to complete the experiential exercises correctly. The data from these five participants, therefore, are not included in the final data set.

### *Setting*

The study was conducted in an experimental cubicle located in the Psychology Department at NUIM. The room was generally free from noise and other potential distractions. After the presentation of a brief verbal outline of the procedure, participants were instructed on the use a stopwatch and audio tape recorder, located on the experimental table. Each participant was then left alone in the room and completed all aspects of the experiment individually. The experimenter remained seated outside the room during this time.

### *Apparatus and Materials*

Participants were presented with a stopwatch to accurately measure the duration of engaging with the experimental exercises (see below). Participants were also provided with a tape-recorder, as an experimental control to ensure all participants successfully completed all aspects of the experiment (see below). In addition to the stopwatch and tape recorder, participants were presented with two written booklets. The first contained a series of psychological assessments and the second contained the basic instructions for completing the experimental sequence.

*Psychological Measures.* All participants completed three self-report psychological measures. The Acceptance and Action Questionnaire (AAQ-49) is designed to measure an individual's level of experiential avoidance (Hayes, Strosahl, et al., 2004, see Appendix F). The Depression, Anxiety and Stress Scales (DASS-21) is designed to measure current levels of depression, anxiety and stress (Lovibond, & Lovibond, 1995, see Appendix G). The Balanced Inventory of Desirable Responding (BIDR) is designed to assess an individual's levels of self-deception and impression management (Paulhus, 1988, see Appendix H). The assignment of participants to one of the nine experimental conditions did not depend upon their questionnaire scores. The measures were simply included to control for any variability across conditions that might account for any differences obtained during subsequent experimental phases.

The *AAQ-49* is a self-report measure of an individual's general levels of emotional avoidance and is similar to the AAQ 35-item version employed in Experiment 1. Individuals are asked to rate the truth of each of 49 statements as it applies to themselves, on a scale of 1 (never true) to 7 (always true). The scales contain 27

negatively reversed items. Total scores range from 49 to 343, with low scores indicating high avoidance and low acceptance, and high scores indicating low avoidance and high acceptance. At the time of the running of Experiment 2, the AAQ-49 suggested superior psychometric properties with respect to the AAQ-35 (see Bond, & Bunce, 2003; Hayes, Strosahl, et al., 2004).

The *DASS-21* is a 21-item inventory that provides a measure of current levels of depression, anxiety and stress. It consists of three sub-scales (Depression, Anxiety and Stress), each of which contains seven questions. Individuals are asked to rate how much each statement applies to themselves over the past week, on a scale of 0 (did not apply to me at all) to 3 (applied to me very much, or most of the time). Total scores for each sub-scale are added together and then doubled (in order to make mean scores comparable with those obtained for the longer 42-item version). Thus total scores for each sub-scale range from 0 (lowest score) to 42 (very severe). The normal level for each of the sub-scales are: Depression < 9; Anxiety < 7; and Stress < 14. The scales possess good psychometric properties with good reliability (alpha values for each of the three sub-scales are: Depression 0.81; Anxiety 0.73; and Stress 0.81) and high internal consistency.

The *BIDR* is used to measure levels of self-deceptive positivity (i.e. providing self-reports that are honest but positively biased) and impression management (i.e. deliberately presenting oneself in a positive light to an audience). The measure consists of two sub-scales -- self-deception and impression management. Scores for each sub-scale range from 0 to 20, with higher scores indicating high levels of self-deception and higher levels of impression management, respectively. The median scores for the self-deception scale are 7.5 ( $SD = 3.2$ ) for males and 6.8 ( $SD = 3.1$ ) for females. The corresponding

scores for the impression management scale are 4.3 ( $SD = 3.2$ ) and 4.9 ( $SD = 3.2$ ) for males and females, respectively. The scales possess good psychometric properties with high levels of internal consistency and test-retest reliability, as well as good concurrent validity with other measures of social desirability.

*Therapeutic scripts.* The scripts employed in the current study were based on those originally used by Masuda et al. (2004), but with some modifications to facilitate experimental improvements, as outlined previously (see below for specific details of the scripts used in the current study).

#### *Experimental Overview*

The current study was a double-blind, independent groups design. Participants were randomly assigned to one of nine experimental conditions, across which Defusion, Thought Control and Placebo rationales and exercises were systematically manipulated -- see Table 4.

Table 4

*An Outline of the Nine Experimental Conditions employed in Experiment 2*

Condition	Rationale	Exercise
<i>Matched Conditions</i>		
Defusion/Defusion	Defusion	Defusion
Thought Control/Thought Control	Thought Control	Thought Control
Placebo/Placebo	Placebo	Placebo
<i>Mismatched Conditions</i>		
Defusion/Thought Control	Defusion	Thought Control
Thought Control/Defusion	Thought Control	Defusion
<i>Unmatched Conditions</i>		
Defusion/Placebo	Defusion	Placebo
Thought Control/Placebo	Thought Control	Placebo
Placebo/Defusion	Placebo	Exercise
Placebo/Thought Control	Placebo	Thought Control

The three *matched* conditions presented participants with two matched components. That is, participants were either presented with: both a defusion rationale paired with a defusion exercise (Defusion/Defusion); a thought control rationale paired with a thought control exercise (Thought Control/Thought Control); or a placebo rationale paired with a placebo exercise (Placebo/Placebo). There were two *mismatched* conditions. Participants in these conditions were presented with two mismatched components. That is, participants were presented with a defusion rationale paired with a thought control exercise (Defusion/Thought Control), or they were presented with a thought control exercise paired with a defusion rationale (Thought Control/Defusion).

The remaining four conditions were *unmatched*. In two of these conditions, participants were presented with a single defusion or thought control rationale which was then paired with a placebo component. Specifically, participants were presented with either a defusion rationale paired with a placebo exercise (Defusion/Placebo), or a



thought control rationale paired with a placebo exercise (Thought Control/Defusion). The remaining two conditions comprised of a defusion or thought control exercise paired with a placebo rationale. Specifically, participants were presented with either a placebo rationale paired a defusion exercise (Placebo/Defusion), or a placebo rationale paired with a thought control exercise (Placebo/Thought Control).

Each of the nine conditions contained the same seven experimental phases across which rationales and exercises were manipulated. Phase 1 involved completing the AAQ-49, the DASS-21 and the BIDR. In Phase 2, participants generated a personally relevant negative statement (e.g. “I am really stupid”), which they were then asked to re-phrase as a single word (e.g. “stupid”). In Phase 3, participants provided Pre-Intervention discomfort, believability and willingness ratings regarding the negative self-relevant thought. The therapeutic rationales were presented in Phase 4. Both the Defusion and Thought Control rationales were presented to participants as enabling them to cope with thoughts and feelings that accompanied the negative self-referential word. In contrast, the Placebo rationale simply involved a short reading about animal communication. The therapeutic exercises were presented in Phase 5 and may or may not have been consistent with the rationale presented in Phase 4 (depending upon whether the condition was matched, mismatched, or unmatched). Participants were instructed to apply the Defusion or Thought Control exercise to the negative word. In contrast, the Placebo exercise simply involved reading about the Gobi desert (Webster, 2002). In Phase 6, participants provided a second series of post-intervention ratings of discomfort, believability and willingness regarding the negative self-relevant thought. Phase 7 contained a series of

short adherence measures designed to determine if participants' reactions to the negative self-referent had been influenced by the rationales and exercises presented to them.

### *Ethical Issues*

In order to conduct the study according to all of the appropriate ethical guidelines as identified by The Psychological Society of Ireland (2003), a series of precautionary measures were explicitly employed in Experiment 2. These were largely consistent with those employed in Experiment 1, with only minor changes of emphasis necessary to facilitate the use of personally relevant statements. Specifically, a greater emphasis here was placed upon informing participants of the possibility of psychological issues arising *directly* from the study. Again, they were emphatically instructed that in the event that this occurred they could contact the thesis supervisor (name provided) and further arrangements, where appropriate, could be made free of charge as a forum for the discussion of these issues with a Chartered Psychologist. At no point during the experiment did any participant withdraw from the study or express dissatisfaction or distress of any kind.

### *Procedure*

*Phase 1: Psychological Measures.* During Phase 1, participants were presented with the first information booklet containing the three psychological measures. The AAQ-49, the DASS-21 and the BIDR were always presented in this order. Participants were instructed to proceed through the booklet as quickly as possible and to avoid spending too much time on each answer. Written instructions at the end of the booklet

then guided participants to Booklet 2 and requested that they set the tape recorder to record.

*Phase 2: Generating a Negative Self-Relevant Thought.* At the beginning of this phase of the experiment participants received the following written instructions:

I would like to begin by asking you to think of *one* particular thought about *you* that regularly enters your mind and that you find very uncomfortable to think about. The thought can be anything. For example, you may have the thought “I am not as good as people think”. The thought just has to be about you and you have to find it very uncomfortable.

Once a relevant thought had been generated, participants were required to write it down on a space provided in the booklet, while simultaneously speaking the words aloud for the purposes of the tape recorder. Participants were then instructed to re-state the thought as a single word. For example, if a participant had the thought “I am stupid”, this could be restated as the single word “stupid”. Once again, participants wrote down the word and spoke it aloud. At the end of Phase 2 (and all subsequent experimental phases), participants were asked to provide a verbal summary aloud of what they had been instructed to do during the phase.

*Phase 3: Pre-Intervention Ratings of the Negative Self-Relevant Thought.* Immediately after Phase 2, participants were asked to rate their levels of discomfort, believability and willingness associated with the negative self-relevant thought they had generated on three separate 11-point Likert scales. Specifically, for discomfort participants were asked: “*How uncomfortable is the thought right now?*” A discomfort rating simply involved circling the number that best reflected the level of discomfort associated with the thought at that particular moment from “0 = Not at all uncomfortable” to “100 = Very uncomfortable” (increasing in 10-point increments). For believability,

participants were asked to: “*Imagine that thought occurred to you right now and, if it occurred, how believable (true) would it be?*” and to rate the believability from “0 = Not at all believable” to “100 = Very believable”. Similarly for willingness, participants were asked to: “*Imagine that thought occurred to you right now and, if it occurred, how willing would you be to think about that thought?*” and rated willingness from “0 = Not at all willing” to “100 = Very willing”.

*Phase 4: Presentation of the Therapeutic Rationale.* Irrespective of the condition to which they had been assigned, all participants received the same initial information about the positive and negative features associated with human language, as follows:

As a species, language (including thoughts and words) gives us both the blessing and the curse of knowledge at the same time. That is, the power of language has pros and cons. There is a light side and a dark side. On the positive side, we can influence the environment and create a comfortable life. Just look around in this room – lights, chairs, central heating and the clothes you are wearing... Without language and our thoughts (e.g. logical thinking), these would not be here. On the dark side, however, we are the only species that worries. And in the most extreme, we are the only species that commits suicide. (It has been suggested that Norwegian lemmings engage in an activity that looks similar to suicide, however, even when a lemming falls into water it will try to climb back out and save itself -- humans do not).

Immediately thereafter, participants were differentiated by condition.

Participants in the three conditions that comprised a Defusion rationale were instructed as follows:

The dark side becomes dominant when we believe our thoughts are literally what they say they are, especially thoughts about ourselves that could be evaluative and judgmental. For example, consider the thought “I am no good”. We tend to think of our thoughts as if they are real. It’s as if you *must be* what your thoughts say you are. For example, you must be no good. But is this really true? Are you really what your thoughts say you are?

What if I told you that thoughts are simply what they are? That is, that thoughts are just thoughts, rather than what they *say* they are. Based on this way of thinking then, maybe you are *not* the person your thoughts tell you that you are.

In contrast, participants in the three conditions that comprised a Thought Control rationale were instructed as follows:

The dark side becomes dominant when we have negative thoughts about ourselves that are evaluative and judgmental. For example, consider the thought “I am no good”. Is this really true? Now consider the thought “I am really good”. This thought is also evaluative and judgmental, but it is a positive thought.

What if I said to you that positive thoughts can be used to take away negative thoughts. That is, when you have the thought “I am a bad person”, you could immediately think instead “I am a good person”. Based on this way of thinking then, maybe you can replace your negative thoughts about yourself with more positive ones.

Participants assigned to the three conditions that comprised a Placebo rationale did not receive instructions for coping with the negative self-referents, but were instead instructed as follows:

Animals operate at a comparatively simple level of interaction and rely heavily on non-verbal gestures to communicate basic needs, for example, food, shelter and signs of danger. Researchers agree that although the language used by humans and more primitive animals have similar features, the two forms are not identical.

Chimpanzees have become the point of reference by which humans compare themselves to all other animals. However, even though chimps share 98.4% of the human genetic code, they are unable to acquire levels of language above those seen in a young child. Chimps can be easily taught up to 120 signs based on American Sign Language, however, unlike humans, they remain unable to generate or produce complex grammatically coherent sentences.

Participants were never made aware of the differences among the experimental conditions nor of the specific condition to which they had been assigned.

*Phase 5: Experiential Exercise and its Application to the Negative Self-Relevant Thought.* At the beginning of Phase 5, all participants were provided with the following brief instructions:

Because it might be a little difficult to understand this, in the next part of the study we will do a little exercise. Okay, the next exercise may seem silly.

At this point, participants were once again divided by condition. Participants presented with the Defusion exercise were instructed as follows:

You will be asked to say a word and then you have to think about what comes to mind. Now say the word “MILK” out loud.

Okay, now say aloud what comes to mind when you say the word “MILK”?

Now say aloud what sort of other things show up when you say “MILK”?

Can you taste it?

Can you feel what it feels like to drink a glass of milk?

Cold, creamy, coats your mouth – right?

Immediately afterwards, these participants were presented with the following information:

What came across your mind were things about *actual* milk and your experience with it. All that happened is that you made a strange sound “MILK” and lots of things showed up. Notice that there isn’t any milk in this room, none at all, but milk was in the room *psychologically*. You were seeing it, tasting it, and even feeling it. And yet, only the word was actually here.

Now, let’s see if you can do an exercise with MILK. It might be a little silly and you might feel embarrassed doing it. What you are asked to do is to say the word “MILK” out loud over and over again, and as rapidly as possible and then notice what happens. Please now set the timer and say the word “MILK” over and over, aloud and as rapidly as possible for approximately 25 *seconds*.

After completing the repetition exercise with MILK, participants were asked to “*Say aloud what came to mind when you kept repeating the word MILK*”. After stating their reactions aloud, participants were then provided with the following instructions:

Did you notice what happened to the milk you thought about before? The creamy stuff goes away. When you said it the first time, it was as if the milk was *actually* here in the room, but all that *really* happened was that you just said the word. The first time you said it, it was psychologically meaningful and it was almost real. But when you said it again and again, did you begin to lose that meaning? Was it as if the words just became a sound, almost like a bird squawking?

After considering the information, participants presented with the Defusion exercise were asked to provide a summary aloud of what they had just been told and what they did. After the summary, the participants were then required to repeat the Defusion exercise with the negative self-relevant thought generated in a previous phase using the following instructions:

What I would now like to suggest to you is that the MILK exercise may be applied to our personal thoughts about ourselves. Isn't it true that thoughts are just thoughts, rather than an actual description of you? Thoughts are just smoke, there isn't anything solid in them. These thoughts are just like the thoughts you had about MILK.

Now your next task is to remember the uncomfortable thought about yourself that you started out with and that you restated as a single word. Could you now say the word aloud and over and over again as rapidly as possible, just like you did with "MILK". Please now set the timer and say your word over and over, aloud and as rapidly as possible for approximately 25 *seconds*.

For example, during the second phase of the experiment a participant may have experienced the thought "I am a stupid person", which was subsequently re-phrased as the single word "stupid". In this case, the participant was required to rapidly repeat this word aloud (e.g. "stupid, stupid, stupid...") for 25secs. At the end of this exercise, participants were once again asked to provide a summary aloud of what they had just done.

In order to present participants with a Thought Control exercise that was formally similar to the Defusion exercise, but that involved a thought control strategy rather than defusion, the MILK exercise was replaced by a black/white exercise in which participants were required to replace the word black with the word white (and positive words associated with it) every time black came to mind. Participants presented with the Thought Control exercise were instructed as follows:

You will be asked to say a negative word and then you have to say the positive word that is its opposite.

Now say the word "BLACK" out loud.

Okay, now say aloud what comes to mind when you say the word "BLACK"?

Now say aloud what sorts of other things show up when you say "BLACK"?

Can you see it?

Can you feel what black feels like?

Dark, thick, hard to see through -- right?

Okay. What came across your mind were things about black.

Okay. Now say aloud the word "WHITE" and notice what sorts of things come to mind when you say the word white.

Can you see it?

Can you feel it?

Bright, shiny, clean -- right?

Immediately after these instructions, participants were presented with the following:

Now, once again say aloud the word 'BLACK' but before you start to think about any of the things that go with it, immediately say the word 'WHITE' and try to think of all the things that go with it instead of 'BLACK'. Please now set the timer for 25 seconds and during this time say the word 'BLACK' aloud first. Then the minute you start to think of what goes with 'BLACK', say the word 'WHITE' and say aloud all of the things that go with the word 'WHITE' instead. Say 'BLACK' as often as it pops into your mind but then say 'WHITE' and try to imagine all the things that go with white rather than black.

Once again, participants were asked to provide a summary aloud of what they had just been told and what they did.

Similar to previous conditions, participants presented with the Thought Control exercise were then asked to apply the same strategy to the negative self-relevant thought generated previously:



What I would now like to suggest to you is that the BLACK exercise may be applied to our personal thoughts about ourselves. Isn't it true that positive thoughts make you feel more comfortable than negative thoughts and are much nicer to have in your mind? These thoughts are just like the thoughts you had about BLACK and WHITE.

Now your next task is to remember the uncomfortable thought about yourself that you started out with and that you restated as a single word. Now, say aloud the word but before you start to think about any of the things that go with it, immediately say an opposite or more positive word and try to think of all the things that go with the positive word instead of the negative one. Please now set the timer for 25 *seconds* and during this time say the negative word aloud first as often as it pops into your mind. Then the minute you start to think of what goes with this, say aloud the opposite positive word and all of the things that go with that word instead.

For example, a participant may have generated the thought "I am a stupid person", which was subsequently re-phrased as "stupid". In this case, the participant was required to state the word "stupid" aloud as often as it came to mind and to replace the negative word "stupid" with more positive words (e.g. "smart, intelligent, funny...") for 25secs. Once again, participants were asked to speak aloud a summary of what they had just done.

The Placebo exercise controlled primarily for time spent during the Defusion and Thought Control exercises and thus involved participants simply reading an extract from an article in the National Geographic about the Gobi Desert (Webster, 2002). The word length of this article was matched as closely as possible to the total number of words presented during the Defusion and Thought Control exercises. At the end of the Placebo exercise, participants were again asked to speak aloud a summary of what they had just done.

*Phase 6: Post-Intervention Rating of the Negative Self-Relevant Thought.* This phase was identical to Phase 3 and required participants to provide (post-rationale and exercise) ratings of discomfort, believability and willingness with regard to the negative self-relevant thought. Again, participants were required to provide a verbal summary of what they did at the end of this phase of the experiment.

*Phase 7: Adherence Measures.* The final phase of the experiment consisted of three 11-point Likert style rating scales that ranged from 0 to 100 (increasing in 10-point increments). The *Honesty Scale* asked participants: “*To what extent did you rate the levels of comfort, believability and willingness honestly?*” (“0 = not at all honest” and “100 = very honest”). The *Difficulty Scale* asked participants: “*Can you tell me how difficult you found it to follow the instructions that were provided for this exercise?*” (“0 = not at all difficult” and “100 = very difficult”). The *Daily Use Scale* presented the following: “*The exercise you have just completed attempted to help you to cope with uncomfortable thoughts. How often do you typically use this approach as a way of coping on a day-to-day basis?*” (“0 = never” and “100 = always”). The Daily Use Scale was not included in conditions that presented a Placebo exercise as no coping exercise was presented in this context. For each of the three scales, participants were required to circle the number that best reflected their level of honesty, the perceived difficulty of instructions and their daily use of the coping strategy with which they had been provided. The completion of Phase 7 marked the end of Experiment 2.

The experiment was completed in a single session and each participant took approximately 35-40mins. to complete the study.

## RESULTS

### *Psychological Measures*

The means and standard deviations for each psychological measure for each condition are presented in Table 5. In general, scores fell within the normal range for

each scale. A series of one-way between groups ANOVAs revealed no significant main effect for condition on any measure (AAQ-49:  $p = .82$ ; DASS subscales: Depression:  $p = .64$ ; Anxiety:  $p = .39$ ; Stress:  $p = .20$ , BIDR subscales: Self-Deception:  $p = .81$  and Impression Management:  $p = .92$ ). Hence, potential differences in subsequent performances could not be attributed to pre-experimental psychological differences between the groups.

Table 5

*Means (and Standard Deviations) for Psychological Measures for Each Condition in*

*Experiment 2*

	AAQ-49	Depression	Anxiety	Stress	Self-Deception	Impression Management
Norms	-	< 9	< 7	< 14	7.50 (3.20)	4.30 (3.20)
Defusion/ Defusion	218.81 (27.60)	8.88 (7.30)	8.25 (7.00)	14.38 (8.71)	4.81 (2.97)	4.88 (3.32)
Defusion / Placebo	222.94 (25.90)	7.06 (8.24)	7.88 (8.47)	14.50 (9.19)	5.94 (3.11)	5.56 (2.87)
Defusion/ Thought Control	218.44 (28.85)	10.13 (10.16)	7.13 (9.35)	15.63 (12.12)	5.19 (3.02)	5.00 (3.50)
Thought Control/ Thought Control	220.69 (25.25)	7.88 (9.08)	7.50 (6.35)	13.50 (7.57)	4.38 (3.10)	5.38 (3.40)
Thought Control/ Placebo	223.19 (34.97)	5.88 (6.34)	8.25 (10.17)	12.38 (8.65)	4.94 (3.04)	6.44 (4.00)
Thought Control/ Defusion	220.44 (37.53)	8.38 (9.39)	5.38 (4.78)	12.13 (8.53)	5.69 (3.98)	5.81 (3.85)
Placebo/ Defusion	224.13 (27.34)	9.88 (7.21)	5.00 (4.73)	13.31 (7.01)	4.25 (2.86)	5.19 (2.81)
Placebo/ Thought Control	231.37 (27.71)	5.44 (4.35)	3.50 (5.49)	8.31 (6.45)	4.38 (3.22)	5.88 (3.18)
Placebo/ Placebo	234.13 (29.87)	6.56 (5.26)	4.44 (3.67)	8.81 (6.40)	5.44 (3.92)	4.81 (3.58)

*Note.* ‘-’ indicates that no normative data is available.

### *Discomfort Ratings*

The mean discomfort ratings at Pre- and Post-Intervention for each condition are presented in Table 6. The majority of participants' scores ranged between 50 and 60 on the scale, where 100 represented maximum discomfort. In all cases, except Placebo/Placebo, discomfort decreased from Pre- to Post-Intervention.

Table 6

*Means (and Standard Deviations) for Discomfort, Believability and Willingness Ratings at Pre- and Post-Intervention in Experiment 2*

Condition	Pre-Intervention	Post-Intervention
<i>Discomfort Ratings</i>		
Defusion/Defusion	67.50 (20.49)	43.75 (24.19)
Defusion/Placebo	55.63 (28.74)	35.63 (28.98)
Defusion/Thought Control	64.38 (21.28)	49.38 (24.07)
Thought Control/Thought Control	58.13 (27.38)	45.63 (25.55)
Thought Control/Placebo	56.88 (26.76)	48.13 (27.86)
Thought Control/Defusion	56.25 (29.41)	43.75 (30.52)
Placebo/Defusion	58.13 (23.73)	31.25 (24.46)
Placebo/Thought Control	51.25 (28.02)	48.13 (24.28)
Placebo/ Placebo	53.75 (24.19)	53.75 (26.55)
<i>Believability Ratings</i>		
Defusion/Defusion	75.00 (15.06)	53.75 (24.18)
Defusion/Placebo	61.25 (23.06)	59.38 (20.16)
Defusion/Thought Control	64.38 (23.08)	50.63 (25.94)
Thought Control/Thought Control	69.38 (28.86)	55.00 (23.66)
Thought Control/Placebo	60.63 (24.62)	55.00 (23.38)
Thought Control/Defusion	56.25 (24.73)	40.00 (27.80)
Placebo/Defusion	75.63 (19.99)	51.88 (28.34)
Placebo/Thought Control	68.75 (26.55)	56.25 (26.55)
Placebo/Placebo	70.00 (20.98)	64.38 (27.32)
<i>Willingness Ratings</i>		
Defusion/Defusion	58.75 (34.42)	60.63 (25.68)
Defusion/Placebo	53.75 (30.74)	47.50 (22.36)
Defusion/Thought Control	63.13 (19.57)	50.00 (27.81)
Thought Control/Thought Control	58.13 (27.62)	55.00 (29.21)
Thought Control/Placebo	48.75 (30.08)	51.88 (29.49)
Thought Control/Defusion	55.00 (23.38)	50.00 (31.83)
Placebo/Defusion	56.25 (26.80)	60.00 (29.21)
Placebo/Thought Control	58.13 (22.28)	56.25 (23.91)
Placebo/Placebo	60.63 (31.93)	56.25 (31.38)

A 9x2 mixed-between-within ANOVA was conducted on the discomfort ratings with condition as the between participant variable and time (Pre- and Post-Intervention) as the within participant variable. Time proved to be significant [ $F(1, 135) = 61.84, p < .001, \eta_p^2 = .31$ ], but condition did not ( $p = .87$ ). However, the interaction effect was also

significant [ $F(8, 135) = 2.98, p = .004, \eta_p^2 = .15$ ]. To further investigate this, nine separate paired-samples t-tests were conducted (one per condition), within a more stringent significance level ( $p = .01$ ) to compensate for the one-tailed prediction of the hypotheses and the multiple comparisons within the dataset. Only four conditions demonstrated a significant decrease in discomfort with large effect sizes:

Placebo/Defusion [ $t(15) = 5.70, p < .001, \eta^2 = .68$ ]; Defusion/Placebo [ $t(15) = 4.57, p < .001, \eta^2 = .58$ ]; Defusion/Defusion [ $t(15) = 4.34, p = .001, \eta^2 = .56$ ]; and Thought Control/Thought Control [ $t(15) = 3.02, p = .009, \eta^2 = .38$ ] -- all remaining  $ps > .01$ .

Defusion, therefore, as matched and unmatched components, significantly decreased discomfort, but Thought Control did so only when *matched*. The lack of effect for Placebo/Placebo ( $p = 1.0$ ) suggested that the observed differences could not be attributed to the passing of time *per se*.

### *Believability Ratings*

The mean believability ratings suggested an overall decrease in levels of believability from Pre- to Post-Intervention (see Table 6). The majority of participants' scores ranged between 50 and 70 on the scale, where 100 represented maximum believability. The 9x2 ANOVA revealed that time was significant [ $F(1, 135) = 65.55, p < .001, \eta_p^2 = .33$ ], but condition was not ( $p = .46$ ). The interaction effect was also significant [ $F(8, 135) = 2.39, p = .02, \eta_p^2 = .13$ ]. A series of paired-samples t-tests indicated significant decreases in believability with large effect sizes for four conditions: Placebo/Defusion [ $t(15) = 4.41, p = .001, \eta^2 = .56$ ]; Defusion/Defusion [ $t(15) = 3.94, p = .001, \eta^2 = .51$ ]; Defusion/Thought Control [ $t(15) = 2.96, p = .01, \eta^2 = .37$ ]; and

Thought Control/Defusion [ $t(15) = 2.93, p = .01, \eta^2 = .36$ ] -- all remaining  $ps > .01$ .

Defusion, therefore, presented as matched, mismatched and as an *unmatched* exercise was associated with significant decreases in believability. Unlike the discomfort ratings, Thought Control was only associated with decreased believability when paired with Defusion. Again, Placebo/Placebo ( $p = .13$ ) had no effect.

### *Willingness Ratings*

The willingness ratings overall showed a pattern of little and unsystematic change (see Table 6). The majority of participants' scores ranged between 50 and 60 on the scale, where 100 represented maximum willingness. A further 9x2 ANOVA revealed that neither time nor condition produced significant main effects, nor did they interact (all  $ps > .19$ ).

### *Adherence Measures*

*Honesty Ratings.* High ratings were generally obtained on the Honesty Scale (for the nine conditions overall:  $M = 87.54, SD = 14.64$ , where 100 represents maximum honesty), and a one-way between groups ANOVA indicated no significant difference across conditions ( $p = .41$ ). In other words, participants were generally honest in reporting their reactions to their negative thoughts and all groups did so with equal honesty.

*Difficulty Ratings.* In general, participants did not appear to experience difficulty following the instructions (nine conditions:  $M = 29.58, SD = 26.65$ , where 100 represents maximum difficulty). However, a one-way between groups ANOVA revealed a



significant main effect for condition [ $F(8, 135) = 3.05, p = .004, \eta_p^2 = .15$ ]. Post-hoc tests indicated that Thought Control/Thought Control generated significantly higher difficulty ratings ( $M = 50.63, SD = 28.63$ ) than: Defusion/Defusion ( $M = 18.13, SD = 25.62, p = .01$ ); Defusion/Placebo ( $M = 18.75, SD = 22.77, p = .01$ ); Placebo/Placebo ( $M = 20.63, SD = 24.07, p = .03$ ); and Placebo/Defusion ( $M = 21.25, SD = 22.17, p = .03$ ). In spite of these differences, combinations between thought control in unmatched and mismatched conditions did not significantly differ from conditions that did not contain a thought control element, thus, it remains possible that the increased difficulty associated with Thought Control/Thought Control resulted from the participants in this group, rather than from the instruction *per se*.

*Daily Use Ratings.* Overall, participants (excluding those in Placebo exercises) did not appear to use the instructions as a daily coping strategy (six conditions:  $M = 21.56, SD = 26.49$ , where 100 represents always using this coping strategy). Interestingly again, a one-way between groups ANOVA revealed a significant main effect for condition [ $F(5, 90) = 4.43, p = .001, \eta_p^2 = .20$ ]. Post-hoc tests indicated that participants in Thought Control/Thought Control were significantly more familiar with the target strategy ( $M = 36.87, SD = 31.14$ ) than both Defusion/Defusion ( $M = 10.0, SD = 12.11, p = .03$ ) and Placebo/Defusion ( $M = 8.13, SD = 14.71, p = .02$ ). Participants in Placebo/Thought Control were also significantly more familiar with their strategy ( $M = 35.0, SD = 25.82$ ) than Placebo/Defusion ( $M = 8.13, SD = 14.71, p = .03$ ). Participants, therefore, were generally more familiar with the use of thought control as a daily coping strategy than with defusion. However, daily use ratings of the strategies presented here were relatively low across all six conditions.

## DISCUSSION

The results for Experiment 2 showed some similarities with the outcomes reported in Experiment 1 and in the research by Masuda et al. (2004). Specifically, all three studies revealed that defusion was associated with significant decreases in the discomfort of negative self-referential thoughts. In comparisons between defusion, thought control and placebo, as well as between instructions and exercises, the current data indicated that the greatest decreases in discomfort were observed when defusion was presented as a *matched* component (i.e. defusion rationale plus exercise) and as an *unmatched* component (i.e. defusion rationale or exercise alone). Overall, Placebo was not associated with decreased discomfort, neither was Thought Control on the whole, with the exception of the matched condition containing both a thought control rationale and exercise.

Defusion also decreased levels of believability associated with the negative self-referent when presented as both rationale and exercise and in almost any combination of components. In this regard, the findings were consistent with Masuda et al.'s (2004) research, but not Experiment 1. In contrast, Thought Control only significantly decreased believability when presented with defusion, but not as an *unmatched* or *matched* component. We recorded no changes in believability for Placebo. To our surprise, Defusion had little or no impact on participants' willingness to experience their negative self-referents. This was inconsistent with Experiment 1, where willingness increased. Masuda et al. did not include measures of willingness, so it was impossible to make comparisons in this regard.

### *Summary of Experiments 1 and 2*

Experiments 1 and 2 investigated the impact of defusion on self-reported changes in discomfort, believability and willingness; attempted to determine the relative utility of defusion exercises versus defusion instructions; and compared defusion with alternative strategies for dealing with negative self-relevant content. Taken together, the results from both studies indicated that defusion exercises significantly decreased the emotional impact of negative self-referents, with reduced effects recorded for defusion instructions.

Experiments 1 and 2 offered a number of advantages over previously published research. For example, Masuda et al. (2004) did not investigate the impact of defusion on willingness and did not use an experimental design that enabled them to determine the relative impact of defusion exercise versus instruction. In addition, each study here incorporated a range of control measures designed to reduce the effect of experimental demand. For example, unlike Masuda et al. where the experimenter delivered the interventions, the experimenter in the current experimental work had minimum contact with participants, with the interventions delivered in an automated format. Furthermore, the experiments presented here also provided a more in-depth look at the impact of defusion on different types of psychological content, as well as providing some insight into the impact of defusion relative to a specific thought control exercise. Taken together therefore, Experiments 1 and 2 appear to make a sound contribution to the very limited existing research database on defusion.

### *The Forthcoming Research*

In spite of its potential contribution to the field, a number of unexpected findings and conflicting evidence across studies raised additional questions about defusion and how it might be investigated. For example, willingness increased in Experiment 1 but not in Experiment 2. In Masuda et al. (2004), defusion instructions alone had no impact on the emotional ratings, but in Experiment 2 here defusion instructions decreased discomfort. A further disparity arose regarding the impact of thought control instructions, with Masuda reporting a significant decrease in believability but Experiment 2 reporting no such effect. Furthermore, in both the Masuda study and in Experiment 2, defusion was associated with decreases in believability but in Experiment 1 here, defusion was associated with increased believability.

Although there were likely a number of specific variables that contributed to these differences (e.g. demand characteristics), it was equally possible that more generic issues pertaining to the self-report measures themselves were an important variable. And indeed, numerous researchers have argued against reliance on direct self-report measures in psychological research, particularly when complex human verbal behaviour is being investigated. In Part II of the doctoral research, we began to investigate this possibility by examining the relationship between implicit measures and defusion.

## **PART II**

## **Chapter 4**

### **Experiments 3, 4 and 5**

**Developing an Implicit Measure of Reactions to  
Positive and Negative Self-Relevant Words  
using the Implicit Relational Assessment  
Procedure (IRAP)**

## Chapter 4

### Experiments 3, 4 and 5

# Developing an Implicit Measure of Reactions to Positive and Negative Self-Relevant Words using the Implicit Relational Assessment Procedure (IRAP)

## INTRODUCTION

Introspective and self-report measures (e.g. questionnaires, interviews, etc.) have traditionally been the centrepiece of clinical research and date as far back as Titchener's work in the early 1900's. The basic assumption is that at some level all individuals have direct and accurate insight into their inner experiences, sensations or thoughts and report these accurately and honestly (Gaillard, Vandenberghe, Destrebecqz, & Cleeremans, 2006). But, there are a number of long-standing generic concerns with these types of assessment tools that are summarised below.

### *(1) Individuals lack insight into their own private events*

Empirical evidence has highlighted discrepancies between individuals' introspective *accounts* of their behaviour and actual behaviour (Nisbett, & Wilson, 1977; Pronin, & Krugler, 2007). In the behavioural tradition, for example, Skinner argued that the discrepancy resulted from the fact that it is inherently difficult for the verbal community to shape how individuals report their inner experiences and

strongly cautioned against the assumption of a causal role of mental events on overt behaviour. In attempts to circumvent these difficulties, behavioural researchers commonly limited experimental response options to yes/no dichotomous responding and simple button presses (Gaillard et al., 2006; Nahmias, 2002).

The concept of experiential avoidance suggests that individuals high in avoidance would have difficulty accessing and describing their own private events, particularly those perceived by the individual as problematic or uncomfortable (Plumb et al., 2004). This may be particularly relevant in attempts to measure defusion, where the targets of defusion are the very experiences that one has chosen to avoid. To overcome these difficulties, clinical researchers have argued for the inclusion of both explicit and behavioural measures in the assessment of psychological avoidance (Gratz et al., 2007; Salters-Pedneault et al., 2004).

## *2) Individuals may be influenced by the demand characteristics of an experiment*

Experimental demand characteristics potentially influence all research interactions with participants, but may be particularly influential on explicit methodologies where individuals interact with, or are required to answer, questions directly. For example, pain researchers have reported that male participants report lower subjective levels of pain in the presence of a female versus male experimenter despite no such differences recorded on heart rate or skin conductance measures (Aslaksen, Myrbakk, Høifødt, & Flaten, 2007).

When presented with a series of measures across a number of experimental phases (e.g. Pre- and Post-Intervention), researchers have also reported attempts by participants to unwittingly appear consistent in their responses across time (Field, & Lawson, 2003). For example, at post-intervention assessments participants may be



able to recall some of their previous responses and try to produce responses of a similar nature. On the contrary, it is equally possible that changes in measures across time reflect the fact that participants have simply forgotten what they were asked previously.

### 3) *Individuals may try to conceal the true nature of their psychological events*

Self-presentation bias is perhaps the most notable weakness of self-report methodologies and potentially affects a range of experimental procedures (Greenwald et al., 2002). For example, participants may simply try to determine the aim of an experiment and thus respond in a manner that is consistent with what they think they are supposed to do, rather than responding on the basis of what they actually experience. Indeed, the '*Hello-Goodbye*' effect is considered to be particularly problematic in clinical settings, when self-report is used to assess patients before and after treatment (de Jong, Passman, Kindt, & van den Hout, 2001).

In what is perhaps a more explicit form of concealment, self-presentation bias reflects attempts to hide an individual's true private events from the public domain (de Jong, 2002; Greenwald, McGhee, & Schwartz, 1998; Teachman, Gregg, & Woody, 2001). This appears to arise from *evaluation apprehension* (e.g. Rosenberg, 1969) or *impression management* (e.g. Leary, & Kowalski, 1990), both of which are particularly relevant in the context of emotionally sensitive material. For example, in studies investigating explicit ratings of the self, many respondents fail to admit to having any negative feelings about themselves at all (Paulhaus, 1991) -- a feature of human nature which seems highly unlikely even from those with strong positive self-regard.

### *Implicit Methodologies and Cognitions*

A number of methodologies have been specifically designed to circumvent the weaknesses associated with explicit measures. These include: affective priming (Fazio, Sanbonmatsu, Powell, & Kardes, 1986); the emotional Stroop Task (Williams, Mathews, & McCleod, 1996); the Go/No-go Association Task (GNAT: Nosek, & Banaji, 2001); the Extrinsic Affective Simon Task (EAST: de Houwer, 2003); the Name Letter Preference Task (Nuttin, 1987); and the Implicit Association Test (IAT: Greenwald et al., 1998).

The range of measures above is commonly referred to as implicit methodologies, based on a working distinction between *explicit* and *implicit cognition*. Explicit cognition is considered to be subjective in nature and is assessed using direct self-report methodologies. In contrast, implicit cognition is thought to be unavailable to self-report and is instead considered to be objective in nature and assessed using indirect or implicit methodologies (Greenwald, 1990). The concept of implicit cognition naturally encompasses the assumption that these events are unconscious and perhaps automatic in nature (Greenwald et al., 1998). Because there is no clear distinction between implicit cognitions and implicit beliefs or attitudes, the term attitude is most commonly used, and Greenwald and Banaji (1995) defined an implicit attitude as:

“introspectively unidentified (or inaccurately identified) traces of past experience that mediate favourable or unfavourable feeling, thought, or action toward social objects” (p.8).

The IAT is undoubtedly the most well established and widely used implicit methodology, with over 250 published articles since 1998 (Greenwald, Nosek, & Sriram, 2006). This measure is considered to have better psychometric properties than any other latency-based measure in terms of split-half reliability, internal consistency

and test-retest reliability (Greenwald, & Nosek, 2001; Schmukle, & Egloff, 2004). The measure also appears to be relatively insensitive to explicit faking (e.g. Banse, Seise, & Zerbes, 2001; Fiedler, & Bluemke, 2005).

The rationale behind the IAT is that it should be easier to map two concepts onto a single response when the concepts are familiar, rather than unfamiliar, in memory. In the seminal IAT study, Greenwald et al. (1998) presented participants with names of flowers (e.g. TULIP), names of insects (e.g. SPIDER), pleasant words (e.g. LOVE) and unpleasant words (e.g. UGLY). They naturally assumed that participants would respond more quickly to *consistent* associations such as pairing the concept “flower” with pleasant words and “insect” with unpleasant words, because both types of associations are consistent with the verbal culture. In contrast, they assumed it would take participants longer to make *inconsistent* associations such as pairing “flowers” with unpleasant words and “insects” with pleasant words. As a latency-based measure, the IAT assesses the strength of the target concept-attribute associations, with the assumption that consistent responding will be superior (i.e. faster) to inconsistent responding – known as the *IAT effect*. As expected, the results from Greenwald et al. supported this assumption.

Of course, the success of the IAT extends much beyond simple associations such as pairing flowers with pleasant words and insects with unpleasant words. Indeed, IAT effects have been reported on a wide range of social and clinical phenomena, including aspects of: developmental psychology (Baron, & Banaji, 2006); neuroscience (Cunningham et al., 2004); and health psychology (Teachman, Gapinski, Brownell, Rawlins, & Jeyaram, 2003) amongst others.

### *Concordance between Implicit and Explicit Attitudes*

There is mixed evidence in terms of published correlations between explicit and implicit measures. Across a series of studies, Nosek and Smyth (2007) measured implicit and explicit preferences for more than 50 evaluative domains and reported implicit-explicit correlations that ranged from weak and positive (<.2) to strong and positive (>.75), with an overall average positive correlation of .48. Other meta-analyses of correlations, however, have reported weaker outputs. For example, Nosek (2005) reported correlations of .36 and others have reported correlations as low as .24 (Hofman, Gawronski, Gschwendner, Le, & Schmitt, 2005).

A number of accounts have been offered in response to these discrepancies.

(1) Some researchers have argued that the discrepancy results from problems associated with explicit measures, especially where the target attitudes are those which individuals would wish to conceal (e.g. Greenwald, & Banaji, 1995). (2) Considerable differences across the two types of methodology have also been highlighted. Specifically, implicit measures are completed under time pressure without opportunity to deliberate (Nosek, 2007). However, when motivation and opportunity to conceal are relatively high, explicit-implicit correlations will be low (Fazio, & Olsen, 2003). Hence, discrepancies are particularly large in the context of socially sensitivity or personally relevant attitudes. (3) Other researchers have argued that the two measures are unrelated and as such are not measuring the same events, but rather reflect the workings of a dual-processing model of human cognition, in which implicit and explicit cognitions are conceptualised as two distinct constructs (Wilson, Lindsey, & Schooler, 2000).

The discordance between explicit and implicit measures has also led researchers to question which, if any, is a better predictor of overt behaviour.

Although there is some evidence that each predicts overt responding to some extent, it has also been suggested that the different measures correlate with different types of behaviour. Specifically, explicit attitudes are considered to predict *deliberate* behaviour, while implicit attitudes predict *spontaneous* behaviour (Fazio, Jackson, Dunton, & Williams, 1995; Gawronski, Hofmann, & Wilbur, 2006; McConnell, & Leibold, 2001).

### *The IAT as a Measure of Implicit Self Regard*

Many of the issues discussed above are pertinent to the potential use of the IAT in clinical research. For example, would clinical interventions have a greater influence over explicit or implicit attitudes and which would be more closely associated with subsequent changes in overt behaviour? The personal and social sensitivities of clinical research also highlight the potential limitations of relying solely on explicit methodologies, but this very fact makes it difficult to determine what size of explicit-implicit correlations one would expect.

One area of personally relevant events that is of clinical significance and that has attracted considerable IAT interest is in the study of attitudes to the self. In one such study of self-esteem, Greenwald and Farnham (2000) assessed automatic associations of the self with positive versus negative attributes and, as expected, reported strong implicit self-positivity for typically-developing undergraduate students (with average to high explicit self-esteem). Interestingly, explicit-implicit correlations here were positive but weak. Bosson, Swann and Pennebaker (2000) provided one of the most extensive investigations into implicit self-esteem. Specifically, the researchers investigated correlations between seven different implicit measures of self-esteem (including the IAT) and a series of well-established explicit

(self-report) measures. Unexpectedly, none of the implicit measures correlated positively with each other and correlated only weakly with the explicit measures.

More recently, Karpinski (2004) used the IAT as a measure of implicit self-esteem in two separate studies. In Study 1, the self was juxtaposed with others, where the latter was either *unspecified* or *specified* as ‘close’ (e.g. best friend, boyfriend, etc.). Unexpectedly, a different pattern of implicit self-esteem emerged across the two contexts of others, with the unspecified other IAT producing significantly higher implicit self-esteem. In Study 2, Karpinski further investigated this effect by altering the valence of the ‘other’ (i.e. the stimuli used as a comparison to the self). Specifically, in one IAT the self was compared to a positive other (Santa Claus) and in another, the self was compared to a negative other (Adolph Hitler). Once again, a different pattern of implicit self-esteem emerged across the two contexts, with higher levels of implicit self-esteem reported in the context of Adolph Hitler. Taken together, the results of these two studies suggest that the content of the stimulus used as a comparison to the self has a considerable influence on implicit self-esteem.

As a result, even a core concept such as self-regard appears to be highly sensitive to context. This outcome highlights the potential difficulties inherent in using implicit methodologies to study personally relevant psychological content that might be of clinical significance.

### *Problems with the IAT*

A number of researchers have pointed to the possibility that some of the difficulties noted above (e.g. weak explicit-implicit correlations and the context susceptibility of implicit attitudes to self) may in fact pertain to the IAT directly as an

implicit measure (Barnes-Holmes et al., 2006). Consider the following quotation from de Houwer (2002):

Greenwald et al. (1998) designed the IAT to assess the strength of associations between concepts in memory. One can argue that beliefs involve more than just associations between concepts. First, beliefs reflect qualified associations. For instance, the belief “I am a bad person” implies *a special type of association* between the concept “self” and the concept “bad”, namely *a directional association* which specifies that “bad” is a property or characteristic of “self”. IAT effects do not reflect *the nature or directionality of an association between concepts*, they can reflect only strength of association. Second, many beliefs involve several associations and several concepts. For instance, conditional beliefs such as “If I do not perform well on a task, then I am an inferior person” involve rather complex structures of qualified associations between several concepts. The IAT cannot be used to directly capture such complex conditional beliefs (also see de Jong et al., 2001, p. 111) . . . In sum, the IAT does not provide a measure of beliefs, nor was it designed to do so. It can only provide an index of associations that are assumed to be involved in certain beliefs and thus indirect evidence for the presence of certain beliefs (pp.117-118, emphases added).

According to de Houwer, therefore, although the IAT may be a robust measure of implicit associations, the procedure does not target the directionality of associations that would be needed for one to fully understand what is being measured. According to Barnes-Holmes et al. (2006), this lack of specificity in the target associations affects all existing implicit methodologies.

#### *The Implicit Relational Assessment Procedure (IRAP)*

The IRAP is a relatively new methodology that primarily attempts to overcome the lack of specificity in the measurement of implicit attitudes, when measured by existing procedures such as the IAT. Although the new procedure bears a strong resemblance to the IAT, it hails from the behavioural tradition, not normally concerned with the study of implicit cognition. However, this all makes more sense when one understands that the IRAP is driven by concepts and procedures that fall under the rubric of Relational Frame Theory (RFT) -- a modern behavioural and functional account of human language and cognition (Hayes et al., 2001). The central

postulate of RFT is that higher-cognitive functioning is composed of relational acts, hence the IRAP explicitly targets stimulus *relations* rather than associations.

The primary difference between the IAT and the IRAP is that in the latter the format of *all* trials is identical. Consider an example from one of the first published studies (Barnes-Holmes et al., 2006). In this IRAP, four words were presented simultaneously on screen. These included: a sample stimulus (the word PLEASANT or UNPLEASANT); a target word easily evaluated as positive or negative (e.g. LOVE or DEATH); and two relational terms (SIMILAR or OPPOSITE) that constituted response options. In order to respond, participants simply selected the response option that best described the relationship between the sample and the target stimuli (e.g. selecting SIMILAR in the combined presence of PLEASANT and LOVE). Similar to the IAT, the IRAP is comprised of consistent and inconsistent trials, where the former contain pre-existing congruent relationships, while the latter do not. In this case, therefore, the researchers predicted that participants would produce shorter response latencies on consistent trials in which they were required to co-ordinate positive words with PLEASANT and negative words with UNPLEASANT (e.g. PLEASANT-LOVE-SIMILAR; PLEASANT-CANCER-OPPOSITE; UNPLEASANT-HOLIDAY-OPPOSITE; and UNPLEASANT-JAIL-SIMILAR). On inconsistent trials, the opposite patterns of responding were required (e.g. PLEASANT-LOVE-OPPOSITE, etc.). Similar to the IAT, the superiority of consistent over inconsistent responding is referred to as the *IRAP effect* and Barnes-Holmes et al. reported several examples of such effects.



### *IRAP and Clinical Phenomena*

Recent empirical research has begun to suggest the utility of the IRAP as a measure of implicit attitudes to self and other clinically relevant phenomena. In Experiment 3 of the original research reported by Barnes-Holmes et al. (2006), the IRAP was used to assess differences in professional attitudes towards normally-developing children and children with Autistic Spectrum Disorder (ASD). Three groups of professionals who differed in their levels of experience with children with ASD were assessed. One group had several *years* experience with children with ASD, another group had *months* of experience and the third group had no experience in this regard at all. All completed both the IRAP and the explicit Challenging Behaviours Attribution Scale (CHABA; Hastings, 1997). During the IRAP, the phrases AUTISTIC SPECTRUM DISORDER and NORMALLY-DEVELOPING were presented as sample stimuli; positive words (e.g. GOOD) and negative words (e.g. DIFFICULT) were the target stimuli; and SIMILAR and OPPOSITE were the response options. In this case, shorter response latencies were predicted on consistent trials in which participants were required to co-ordinate positive words with normally-developing and negative words with ASD. On inconsistent trials, the opposite patterns of responding were required (i.e. co-ordinate negative words with normally-developing children and positive words with ASD).

The results of this experiment indicated that all three groups showed an implicit bias in favour of normally-developing children and against children with ASD on the IRAP and no differences were recorded between the groups on this measure. In contrast, the groups differed significantly on the CHABA, with the most experienced group showing significantly more positive attitudes to children with ASD

than the least experienced. Based on this discrepancy therefore, it was hardly surprising that the researchers reported poor explicit-implicit correlations.

In the only publicly available research to date using the IRAP as a measure of implicit self-regard, Vahey, Barnes-Holmes, Barnes-Holmes and Stewart (in press) compared the IRAP performances of undergraduate students relative to two groups of prisoners. The latter groups were divided according to the part of the prison in which they currently resided (i.e. Open Area versus Main Block). It was predicted that these three groups would differ in relation to levels of self-esteem, with undergraduate students categorised as having normal to high self-regard; prisoners in the Open Area categorised as having self-regard that was probably lower than undergraduates but higher than prisoners in the Main Block; and prisoners in the Main Block categorised as having the lowest self-regard. All participants completed both the IRAP and the explicit Rosenberg Self-Esteem Scales (RSES; Rosenberg, 1965). During the IRAP, the words SIMILAR and OPPOSITE were presented as sample stimuli; positive words (e.g. GOOD) and negative words (e.g. BAD) were the target stimuli; and PARTICIPANT'S NAME and NOT PARTICIPANT'S NAME were the response options. In this case, shorter response latencies were predicted on consistent trials in which participants were required to co-ordinate positive words as similar to their own name and negative words as similar to the other response option (i.e. not participant's name). On inconsistent trials, the opposite patterns of responding were required (i.e. co-ordinate negative words with their own name and positive words with 'not their own name').

The results indicated that, as predicted, the three groups of participants were differentiated according to explicit self-esteem with undergraduates reporting the highest levels of explicit self-esteem on the RSES and prisoners in the Main Block the

lowest. Interestingly however, the IRAP differentiated the Privileged Area prisoners and undergraduates from the Main Block area, but did not differentiate the former from each other. Moderate positive correlations were reported between the implicit and explicit measures. These findings go some way towards suggesting the possible utility of the IRAP as a measure of socially and personally sensitive implicit attitudes.

### *The Current Research (Experiments 3, 4 and 5)*

The basic aim of this part of the research program was to begin to incorporate the use of implicit measures into the experimental study of defusion, in an attempt to supplement the previous work that had relied heavily on self-report measures. However, given the complexities of the concept of defusion itself, the fact that it had only a very limited history as the subject of experimental scrutiny and that there was no history of using the IRAP in this context, the three studies reported in the current chapter were exploratory. Their primary aim was to begin to create an IRAP that would tap into implicit self-relevant content in a manner that would subsequently allow us to manipulate these implicit attitudes through defusion. In other words, we wanted to construct an IRAP where self-esteem might be low, thus allowing for improvement through a defusion intervention.

In all cases, participants were undergraduate students who presented with average to high self-regard. Across the three studies, the primary IRAP manipulation centred on the sample stimuli. In Experiment 3, the samples were I AM versus I AM NOT; in Experiment 4 the samples were I AM versus OTHERS ARE; and in Experiment 5 the samples were IAM versus I SHOULD BE. All participants were also presented with a range of relevant explicit measures.

## EXPERIMENT 3

### METHOD

#### *Participants*

Sixteen participants (all female) aged between 18 and 21 years ( $M = 20$ ) comprised the final sample of Experiment 3. All were undergraduate students, recruited through Faculty announcements at NUIM. All had a high level of fluency in English and had normal, or corrected-to-normal, vision. No incentives of any kind were offered for participation. In addition to the 16 participants who completed the experiment, three participants failed to reach the required criteria in either the training or test phases of the IRAP, hence their data was excluded from further analyses.

#### *Setting*

Experiment 3 was conducted in the Computer Laboratory of the Psychology Department at NUIM. Noise and distraction were kept to a minimum throughout and the experimenter remained in the room during all aspects of the study. During IRAP practice blocks, the Experimenter remained in close proximity to participants, but moved to a more distant location during other parts of the procedure. Several participants may have been present in the Laboratory at the same time, but were positioned some distance apart and no two commenced at the same time.

#### *Apparatus and Materials*

*Psychological Measures.* Participants were required to complete three psychological measures, including the Depression Anxiety Stress Scale (DASS-21; Lovibond, & Lovibond, 1995 -- see Appendix G and Chapter 3); the Rosenberg Self-

Esteem Scale (RSES; Rosenberg, 1965 -- see Appendix I); and the Acceptance and Action Questionnaire-II (AAQ-II; Bond et al., 2008 -- see Appendix J).

The *DASS-21* is comprised of three self-report scales designed to measure negative emotional states associated with depression, anxiety and stress (see Chapter 3).

The *RSES* is one of the most widely used self-report measures of global self-esteem. It is a 10-item scale on which participants rate their agreement with each statement along four points, ranging from “1” (strongly agree) to “4” (strongly disagree). The scales contain five negatively reversed items. Total scores range from 10 to 40, with higher scores representing higher levels of self-esteem and lower scores representing lower levels of self-esteem. The RSES generally has high reliability test-retest correlations are typically in the range of .82 to .88 (Blascovich, & Tomaka, 19993) and Cronbach's alpha value of .88 (Greenberger, Chen, Dmitrieva, & Farruggia, 2003).

The *AAQ-II* is a revised version of the AAQ-49, often referred to generically as a measure of experiential avoidance and psychological flexibility. The measure consists of 10 items on which participants rate the level of truth associated with each statement on a 7-point scale, ranging from “1” (never true) to “7” (always true). The scale contains seven negatively reversed items. Total scores range from 10 to 70, with high scores representing high levels of acceptance/low levels of avoidance and low scores representing low levels of acceptance/high levels of avoidance. The mean scores for a sample of university students is 50.72 ( $SD = 9.19$ ). The alpha value of the AAQ-II is strong at .83, while the scales also possess good test-retest reliability: .80 (at 3 months) and .78 (at 12 months -- Bond et al., 2008).

*IRAP-Related Response Scales.* Participants also completed four IRAP-related Response Scales, designed to assess their explicit attitudes to the self-relevant stimuli that would be presented during the IRAP. All four Likert rating scales, therefore, measured participants' explicit attitudes to the twelve target words -- six positive (KIND, CONFIDENT, TRUSTING, HONEST, SECURE and POPULAR) and six negative words (SELFISH, SELF-CONSCIOUS, JEALOUS, FAKE, INSECURE and LONELY -- see Appendix K). The negative words were selected from Experiment 2 as the six most frequently occurring uncomfortable words that had been generated by participants. The positive words were simply selected as opposites. One of the four scales assessed participants' *feelings towards* each of the target words; a second assessed levels of *discomfort* associated with the words; a third assessed levels of *believability* associated with the words; and the fourth assessed levels of *willingness* in terms of thinking about the words. Each scale ranged from 1 (extremely negative, extremely uncomfortable, extremely believable and extremely unwilling) to 7 (extremely positive, extremely comfortable, extremely unbelievable and extremely willing). Participants indicated their preferred choice by circling the corresponding number on each of the individual rating scales.

*IRAP.* Participants completed all IRAP trials on a Dell personal computer with a colour monitor, Intel Pentium 4 processor and Windows XP operating system. The computer program was written in Microsoft Visual Basic 6.0 and presented all instructions and stimuli relating to the IRAP trials, and recorded participant responses.

The stimuli presented by the IRAP in Experiment 3 consisted entirely of groups of words. The *sample stimuli* consisted of two phrases: I AM and I AM NOT. The *target stimuli* were all single nouns. Six had positive connotations (KIND, CONFIDENT, TRUSTING, HONEST, SECURE and POPULAR) and six had

negative connotations (SELFISH, SELF-CONSCIOUS, JEALOUS, FAKE, INSECURE and LONELY). The two *response options* consisted of the relational terms TRUE and FALSE. The stimulus arrangements for the IRAP are presented in Table 7.

Table 7

*The Stimulus Arrangements Employed in Experiment 3*

<b>Sample Stimulus 1</b> I AM	<b>Sample Stimulus 2</b> I AM NOT
<b>Response Option consistent with Sample 1</b> TRUE	<b>Response Option consistent with Sample 2</b> FALSE
<b>Target Stimuli consistent with Sample 1</b>	<b>Target Stimuli consistent with Sample 2</b>
KIND CONFIDENT TRUSTING HONEST SECURE POPULAR	SELFISH SELF-CONSCIOUS JEALOUS FAKE INSECURE LONELY

*Experimental Overview*

Each participant completed all three experimental phases in the same order. During *Phase 1*, they completed the three psychological measures (the DASS-21, the RSES and the AAQ-II), as well as the four IRAP-related Response Scales. *Phase 2* comprised of the IRAP. In *Phase 3*, participants received a second exposure to the IRAP-related Response Scales.

*Ethical Issues*

In order to conduct the study according to all of the appropriate ethical guidelines as identified by The Psychological Society of Ireland (2003), a number of specific measures were put in place. These were consistent with those outlined

previously for Experiment 1, with only minor adjustments necessary to facilitate the change in methodologies to tailor instructions to the current experiment.

### *Procedure*

*Phase 1: Pre-Experimental Measures.* Participants were individually presented with two questionnaire booklets in written format. The first booklet contained the DASS-21, the RSES and the AAQ-II. The second booklet contained the four IRAP-Related Response Scales.

*Phase 2: The IRAP.* The IRAP program commenced with a series of automated ethical guidelines and instructions as follows:

Our research investigates cognitive processes that are used in decisions that involve memory. We are seeking to develop and test theories of cognitive processes that occur inside and outside of awareness in the routine use of memory.

Stimuli will be presented on this display screen and your responses will be entered on the keyboard.

The research assumes that you can read English fluently and that your vision is normal or corrected-to-normal. If you do not consider yourself fluent in English, or if your vision is not normal or corrected-to-normal and **ESPECIALLY IF YOU ARE HAVING SOME DIFFICULTY READING THIS DESCRIPTION**, please ask the Experimenter now whether or not you should continue.

Your identity as a participant is confidential. Further, you are free to discontinue participation at any time, without penalty. In keeping with standard practice, your data may be retained for 5 years or so, during which time only the investigators on this or successor projects will have access to them.

**PLEASE NOW READ THE STATEMENT BELOW, WHERE YOU WILL BE ASKED TO RESPOND TO A STANDARD INFORMED CONSENT QUESTION.**

I have read the description of the procedure. I understand that the questions I may have about this research will be answered by Professor Barnes-Holmes or one of the other researchers working on this project.

If you consent to participate in the research that has been described on the preceding display pages you should now read the Instructions for the sorting tasks below.

[INSTRUCTION: If you wish to ask any questions first, alert the experimenter now. **IF YOU WISH NOT TO PROCEED**, you should inform the experimenter].



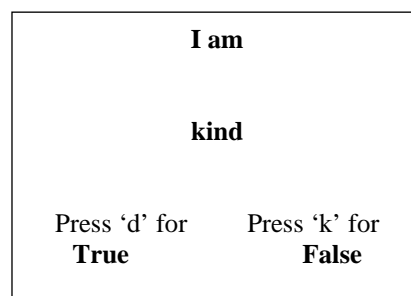
Once participants had read this information and consented to continue with the experiment, further instructions for task completion were presented with illustrated on-screen examples as follows:

Shown below are illustrations of the four different types of task that will be presented repeatedly in this part of the experiment. To help you understand the tasks, each of the four illustrations is explained immediately underneath. Please examine each illustration and then read carefully the explanation attached to it. Please make sure that you understand each task before continuing with the experiment.

**IMPORTANT: From trial to trial the positioning of the response options (True and False) will vary randomly between left and right. NOTE: During the experiment a range of other words apart from “Kind” and “Selfish” will also be presented.**

**REMEMBER: From trial to trial the positioning of the response options (True and False) will vary randomly between left and right.**

For illustrative purposes, one (of a possible four) on-screen examples is presented in Figure 3.



*Figure 3.* An example of the on-screen illustrations presented to participants in Experiment 3.

An explanation accompanied each of the four illustrations in order to ensure that participants understood the appropriate responses that were required. For example, the explanation for the illustration shown in Figure 3 read as follows:

If you select “True” by pressing the ‘D’ key, you are stating that “I am kind.”  
If you select “False” by pressing the ‘K’ key, you are stating that “I am NOT kind.”

Thus, participants were required to determine whether the relationship between I AM and KIND is TRUE or FALSE and to select the correct response key accordingly (‘d’

for TRUE and 'k' for FALSE). Immediately after this, the following automated instructions were presented:

During the experiment you will be asked to respond as **quickly and accurately** as you can across all trials.

The relating tasks will be presented in short *sessions* that are separated by the appearance of instructions on the computer screen. You can take a short break if you like while the instructions are on on-screen.

During each short session the relating task follows one general rule. An incorrect response on any trial is signalled by the appearance of a red 'X' in the centre of the screen. To remove the red 'X' and move on to the next trial please press the correct response key quickly.

After each session, further instructions will appear and they will tell you that the general rule that applied in the previous session is now completely reversed. Please pay close attention to these instructions and do your best to follow them.

So, just to clarify, there will be only *two* general relating rules and so the first thing you should do at the beginning of each session is to discover the rule by using the feedback you get in the form of the red 'X'. It is very important to understand that sometimes you will be required to respond to the tasks in a way that *agrees* with what you believe and at other times you will be required to respond in a way that *disagrees* with what you believe. **This is part of the experiment.**

The first two sessions are for *practice* only and these are repeated until you respond accurately on at least 80% of the relating trials and respond faster, on average, than 3000 milliseconds (i.e. 3 seconds). When you complete the practice phase, the *test-phase* will then start. Remember, you should try to make your responses as **accurately and quickly** as possible.

This point in the procedure marked the beginning of the automated practice IRAP trials. Each trial was presented as a single screen and the format of all trials was identical. On each trial, one sample stimulus, one target stimulus and the two response options were presented simultaneously on-screen. The sample (I AM or I AM NOT) appeared at the top of the screen, with the target stimulus in the centre below and the two relational terms at the bottom left- and right-hand corners (see Figure 3). On all trials, the response option on the left was accompanied by the instruction "Press 'd' for" and the response option on the right by the instruction "Press 'k' for", so that the location of each matched the relevant keys on the left- and right-hand side of the keyboard. Although the response instructions always remained in the same locations

on screen in order to match the keys on the keyboard, the relational terms (TRUE and FALSE) alternated in location between the trials in a random order (i.e. sometimes 'd' was associated with TRUE, other times with FALSE).

*Feedback.* Throughout the IRAP, feedback on incorrect responding was presented in the form of a red 'X' that appeared in the centre of the screen under the target word. Participants were prevented from continuing to the next trial until they provided a correct response and the 'X' remained on-screen until this occurred. Correct responding was not consequated by feedback, but was followed by a clear screen for 400ms. before the next trial appeared.

*Consistent versus Inconsistent Responding.* The IRAP systematically juxtaposes consistent versus inconsistent trials. In the case of this IRAP, trials denoted as *consistent* comprised of associating the self (I AM) with positive words and not-the-self (I AM NOT) with negative words. During consistent trials, therefore, correct responding involved the following, for example: I AM-HONEST-TRUE; I AM-SELFISH-FALSE; I AM NOT-SELFISH-TRUE; I AM NOT-HONEST-FALSE and incorrect responding involved I AM-SELFISH-TRUE, etc. In contrast, during the *inconsistent* trials, correct responding involved, for example, I AM-HONEST-FALSE; I AM-SELFISH-TRUE; I AM NOT-SELFISH-FALSE; I AM NOT-HONEST-TRUE and incorrect responding involved I AM-SELFISH-FALSE, etc.

*Trial-types.* The current IRAP consisted of four basic trial-types: I AM-Positive; I AM-Negative; I AM NOT-Positive; and I AM NOT-Negative (see Figure 4). Within each block of 24 trials, the sample I AM was presented on 12 trials with I AM NOT presented on the other 12. The presentation of sample stimuli was randomised throughout each block. In the presence of each sample, each of the twelve

target words (six positive and six negative) appeared twice. The locations of TRUE and FALSE were randomised across trials.

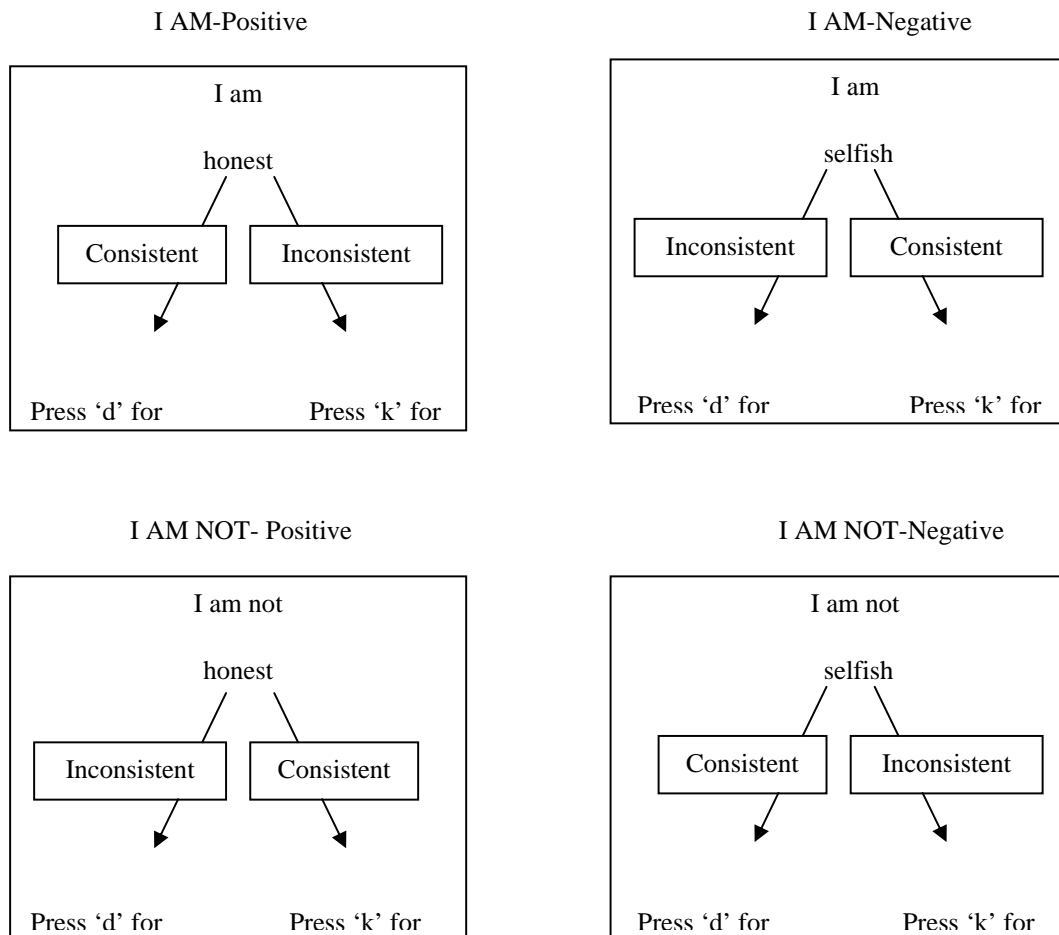


Figure 4. Examples of the four IRAP trial-types used in Experiment 3. The superimposed arrows with text boxes indicate the responses deemed consistent or inconsistent, but these boxes and arrows did not appear on screen during the experiment.

*IRAP Sequence.* The IRAP sequence consisted of a minimum of two practice blocks always followed by six test blocks. Practice and test blocks were identical in format but the short instruction that preceded a practice block was “*This is a practice - errors are expected*”, while the instruction “*This is a test - go fast, making a few errors is ok*” preceded test blocks. Each block consisted of either

consistent or inconsistent trials *only* and the sequence always comprised of alternating blocks (consistent then inconsistent or vice versa). Thus, if a participant was exposed to a consistent block first, the next block was inconsistent, followed by a consistent block, and so on. In the current experiment, half of the participants were exposed to a consistent block first (referred to as *consistent-first*) and the other half were exposed to an inconsistent block first (referred to as *inconsistent-first*).

*Practice Trials.* All participants completed a minimum of two practice blocks (one consistent and one inconsistent). They were required to achieve an accuracy criterion of 80% correct and a response latency of 3000ms. -- failure to do so meant that the participant was immediately exposed to a second pair of practice blocks. Failure across four pairs of practice blocks resulted in termination of participation in the study. However, if criterion during the practice blocks was reached, participants continued to the test blocks. There were six test blocks in total (three of which were consistent trials and three were inconsistent trials). Participants were not made explicitly aware of the alternating block sequence or the terms “consistent” and “inconsistent”. They were, however, provided with a brief instruction at the beginning of each block which emphasised that the responding required for the pending block was opposite to that from the previous block:

Important. During the next phase, the previously correct and incorrect answers are reversed. This is part of the experiment. Please try to make as few errors as possible. In other words, avoid the red X.

At the end of each block of trials, participants were presented with a summary of their performances during the previous block in terms of accuracy percentage and mean response latency. Participants were permitted to pause between blocks and pressed the space bar when they were ready to proceed to the next block. Following completion of the last test block, participants were presented with the following

instruction: “*Thank you. This is the end of the sorting task. Please report to the experimenter.*”

*Phase 3: Post-experimental Measures.* Following completion of the IRAP, participants were presented with another questionnaire booklet in written format. This contained four IRAP-related Response Scales and was identical to that presented previously in Phase 1.

Participants were thanked for their co-operation and time. They were then appropriately debriefed about the research and invited to ask any questions. Although made available to them, no participants opted for short breaks at any point. All participants completed the study in a single experimental session that lasted between 20 and 30mins. in total.

## RESULTS

### *IRAP Analysis*

The IRAP program automatically records accuracy levels and response latencies for all trials. *Accuracy* was defined in terms of the first response emitted on a trial (and thus incorporates a correction procedure for incorrect responding). Accuracy data were not incorporated into the analyses but were employed as a screening procedure to exclude participants who failed to achieve a criterion of 80% accuracy at any point. The data presented here represents sixteen females who successfully met the criteria outlined above.

*D<sub>IRAP</sub>-algorithm Data.* The primary datum used in the current analyses was response latency defined as the time in milliseconds that elapsed between the onset of the trial and a correct response emitted by a participant. The response latency data for

each participant were transformed into  $D_{IRAP}$  scores (Cullen, & Barnes-Holmes, in press) using an *adaptation* of the Greenwald, Nosek and Banaji (2003) D-algorithm. The steps involved in calculating the  $D_{IRAP}$  scores were as follows: (1) only response latency data from the six test blocks were used; (2) latencies above 10,000ms. were removed from the dataset; (3) if the data from a participant contained more than 10% of test block trials with latencies less than 300ms. that participant was removed from the analyses; (4) twelve standard deviations for the four trial-types were calculated: four for the response-latencies from test blocks 1 and 2, four from the latencies from test blocks 3 and 4 and a further four from test blocks 5 and 6; (5) 24 mean latencies were then calculated for the four trial-types in each test block; (6) difference scores for each of the four trial-types were calculated, for each pair of test blocks; (7) each difference score was then divided by its corresponding standard deviation from step 4, yielding 12  $D_{IRAP}$  scores -- one score for each trial-type for each pair of test blocks; (8) four overall trial-type  $D_{IRAP}$  scores were calculated by averaging the three scores for each trial-type across the three pairs of test blocks; (9) two  $D_{IRAP}$  scores, one for *Pro-Me* and one for *Anti-Me*, were then calculated by averaging the two I AM and then the two I AM NOT trial-type scores, respectively; (10) an overall  $D_{IRAP}$  score was also calculated by averaging all 12 trial-type scores from step 7.

The  $D_{IRAP}$ -algorithm can be used without regard to trial-types to calculate a single  $D_{IRAP}$  statistic for each participant as above, or alternatively it can be used with data segregated according to trial-type to produce two  $D_{IRAP}$  statistics for each participant -- one for Pro-Me trial-types and one for Anti-Me trial-types. The section below contains analyses involving  $D_{IRAP-(trial-type)}$  data, rather than a single D score. The latter, however, will be employed in subsequent comparisons between the implicit and explicit measures.

Figure 5 presents the mean  $D_{IRAP}$  scores on the pairs referred to as Pro-Me (I AM-Positive and I AM-Negative trial-types combined) and Anti-Me (I AM NOT-Positive and I AM NOT-Negative). The data indicated that participants showed a strong implicit bias towards the self, with little or no opinion with regards to what they are not.

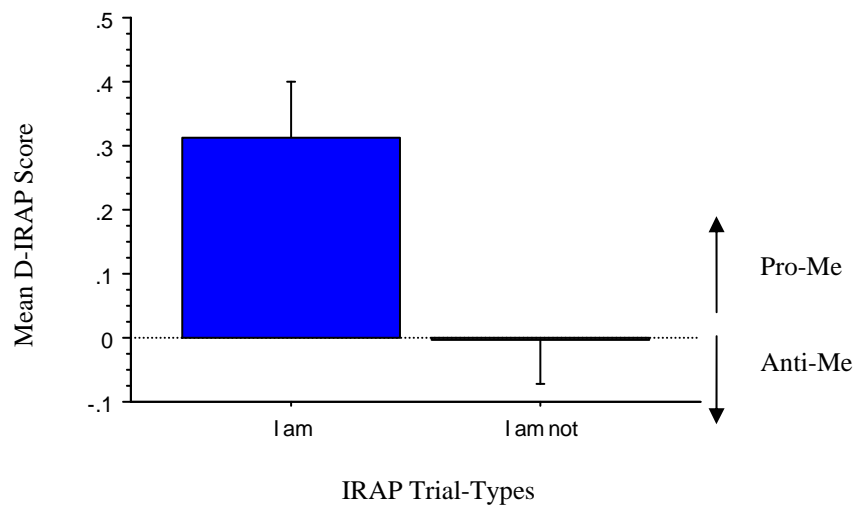


Figure 5. The mean  $D_{IRAP-(trial-type)}$  scores and standard error bars for the sample stimuli in Experiment 3.

A one-way repeated measures ANOVA revealed a large significant main effect for trial-type [ $F(1, 15) = 17.04, p = .0009, \eta_p^2 = .53$ ], indicating a difference between the Pro-Me and the Anti-Me trial-types. Two one sample t-tests indicated that the I AM trial-types differed significantly from zero,  $t(31) = 3.68, p = .0009, \eta^2 = .47$ , but as expected I AM NOT was non-significant ( $p = .95$ ). Overall therefore, the IRAP indicated that participants had strong positive implicit self-regard, but had no strong opinions with regard to what they were not. Put simply, the positive self-regard did not imply that they had to confirm that they were not negative.

#### *Explicit Measures*



*Psychological Measures.* The means scores for each of the DASS-21 subscales fell within the normal range: Depression ( $M = 7.88$ ,  $SD = 5.44$ ; Norm  $< 9$ ); Anxiety ( $M = 6.38$ ,  $SD = 7.49$ ; Norm  $< 7$ ); Stress ( $M = 13.75$ ,  $SD = 9.66$ ; Norm  $< 14$ ). As expected, participants also produced high RSES scores ( $M = 21.88$ ,  $SD = 4.76$ ), indicating high levels of self-esteem, as well as high AAQ-II scores ( $M = 49.75$ ,  $SD = 9.0$ ; Norm = 50.72 ( $SD = 9.19$ ), indicating high levels of psychological flexibility (high acceptance/low avoidance).

*Four IRAP-Related Response Scales.* Mean scores (and standard deviations) for Pre- and Post-IRAP ratings for the positive and negative IRAP words are presented in Table 8. Participants overall rated the negative words as more negative, more uncomfortable and more unbelievable (i.e. less believable), and they were more unwilling to experience them, relative to the positive words. Although there were some apparent changes on these scales from Pre- to Post-IRAP, these appeared to be unsystematic.

Table 8

*Mean Ratings (and Standard Deviations) for the Positive and Negative IRAP Words at Pre- and Post-IRAP in Experiment 3*

IRAP-Related Response Scale	Pre-IRAP	Post-IRAP
<i>Feeling</i>		
Positive	5.90 (1.27)	6.20 (0.91)
Negative	2.53 (1.47)	2.12 (0.93)
<i>Discomfort</i>		
Positive	6.34 (0.95)	6.00 (1.28)
Negative	2.36 (1.19)	2.74 (1.71)
<i>Believability</i>		
Positive	2.92 (1.51)	2.93 (1.69)
Negative	4.69 (1.73)	4.65 (1.78)
<i>Willingness</i>		
Positive	6.01 (1.38)	5.97 (1.48)
Negative	4.15 (1.59)	3.38 (1.84)

*Note* .Higher scores indicate words were rated as more positive, more comfortable, less believable and participants were more willing to experience them.

Four separate 2x2 repeated measures ANOVAs were conducted (one for each of the IRAP-related Likert Scales) with time (Pre-IRAP and Post-IRAP) and word type (positive and negative) as the within participant variables. The results from the *Feeling* Scale indicated a large significant main effect for word type [ $F(1, 30) = 247.28, p < .0001, \eta_p^2 = .89$ ], but not for time ( $p = .67$ ) and the interaction was also non-significant ( $p = .15$ ). The results were identical for the *Discomfort* Scale -- word type [ $F(1, 30) = 166.94, p < .0001, \eta_p^2 = .85$ ], time and interaction effect ( $p$ 's = .46 and .06, respectively). They were also the same for the *Believability* Scale -- word type [ $F(1, 30) = 15.47, p = .0005, \eta_p^2 = .34$ ], time and interaction ( $p$ 's = .94 and .10, respectively). And again, the results were the same for the *Willingness* Scale -- word type [ $F(1, 30) = 69.53, p < .0001, \eta_p^2 = .70$ ], time and interaction ( $p$ 's = .32 and .19, respectively).

Taken together, these results suggest that participants rated the positive words as significantly more positive, more comfortable, more believable and easier to experience than the negative words. No differences were recorded between Pre- and Post-IRAP ratings on any of the four scales.

### *Correlations between Implicit and Explicit Measures*

*Psychological Measures.* Correlational analyses were conducted between the three implicit measures (IRAP-Overall score, IRAP-I AM score and IRAP-I AM NOT score) and each of the psychological measures (depression, anxiety, stress, RSES and AAQ-II). The results indicated very weak correlations between the psychological measures and the IRAP scores (all  $r$ 's < .25).

*Four IRAP-Related Response Scales.* Correlational analyses were conducted between the implicit measures (IRAP-Overall score, IRAP-I AM score and IRAP-I AM NOT score) and each of the four explicit IRAP-related Response Scales (Feeling, Discomfort, Believability and Willingness). Results indicated a weak pattern of correlation between the measures (all  $r$ 's < .27).

## DISCUSSION

The results from Experiment 3 indicated that participants had a strong positive implicit self-regard, but no strong opinions as to what they were not – thus indicating that their positive self-regard was dependent on confirming that they were positive and not negative, but was not dependent on disconfirming that they were negative and not positive. The results of the explicit measures indicated that participants overall scored within the normal range on each of the psychological measures. As expected,

participants also rated the positive words as more positive, more comfortable, more believable and easier to experience than the negative words. Furthermore, these ratings did not change from Pre- to Post-IRAP. Finally, there were no correlations between the implicit measures and any of the explicit measures (psychological measures and IRAP-related Response Scales).

Although the results from the IRAP indicated that participants had a strong positive implicit self-regard, the verbal relation employed here (i.e. I AM versus I AM NOT) is only one of many that an individual may use to determine their self concept. It remained possible, therefore, that different results would be obtained if participants were asked to compare the self to other individuals, for example -- this was the focus of Experiment 4.

## EXPERIMENT 4

### METHOD

#### *Participants*

Sixteen participants (all female) aged between 18 and 26 years ( $M = 22$ ) comprised the final sample of Experiment 4. All were undergraduate students, recruited through Faculty announcements at NUIM. All participants had a high level of fluency in English and normal, or corrected-to-normal, vision. There were no incentives for participation. All participants recruited for this experiment successfully completed all aspects of the experiment.

#### *Setting*

All aspects of the setting for Experiment 4 were identical to the previous experiment.

#### *Materials and Apparatus*

*Explicit Measures.* The three psychological measures and the four IRAP-related Response Scales used previously were employed again in Experiment 4.

*IRAP.* The primary difference between the IRAP employed in Experiments 3 and 4 is that the current experiment presented I AM and OTHERS ARE as sample stimuli. The target stimuli and response options remained the same. Naturally, this also necessitated some changes to the instructions provided to participants. The stimulus arrangements employed in Experiment 4 are presented in Table 9.

Table 9

*The Stimulus Arrangements Employed in Experiment 4*

<b>Sample Stimulus 1</b> I AM	<b>Sample Stimulus 2</b> OTHERS ARE
<b>Response Option consistent with Sample 1</b> TRUE	<b>Response Option consistent with Sample 2</b> FALSE
<b>Target Stimuli consistent with Sample 1</b> KIND CONFIDENT TRUSTING HONEST SECURE POPULAR	<b>Target Stimuli consistent with Sample 2</b> SELFISH SELF-CONSCIOUS JEALOUS FAKE INSECURE LONELY

*Procedure*

All procedural aspects and ethical guidelines of Experiment 4 were identical to the previous study.

*Phase 1: Pre-Experimental Measures.* Participants first completed the DASS-21, the RSES and the AAQ-II, as well as the four IRAP-related Response Scales.

*Phase 2: The IRAP.* The IRAP for Experiment 4 consisted of four basic trial-types: I AM–Positive; I AM–Negative; OTHERS ARE–Positive; and OTHERS ARE–Negative (see Figure 6). Within each block of 24 trials, the sample stimulus I AM was presented on 12 trials, with OTHERS ARE presented on the other 12 trials. The presentation of sample stimuli was randomised throughout each block. In the presence of each sample, each of the twelve target words (six positive and six negative) appeared twice. The locations of TRUE and FALSE were randomised across trials.

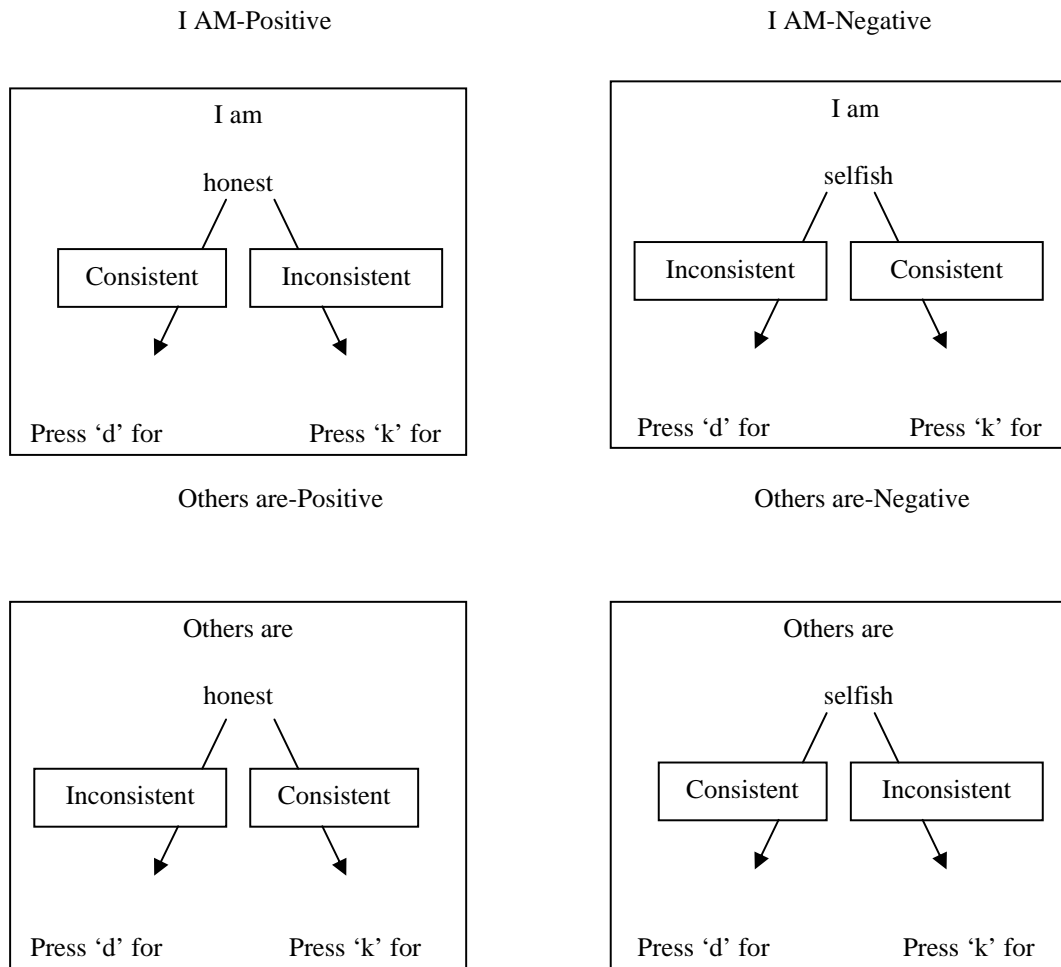


Figure 6. Examples of the four IRAP trial-types from Experiment 4. The superimposed arrows with text boxes indicate the responses deemed consistent or inconsistent, but these boxes and arrows did not appear on screen during the experiment.

*Consistent versus Inconsistent Responding.* In the current IRAP, correct responses on *consistent* trials involved associating I AM with positive words and OTHERS ARE with negative words (e.g. I AM-HONEST-TRUE; I AM-SELFISH-FALSE; OTHERS ARE-SELFISH-TRUE; OTHERS ARE-HONEST-FALSE and incorrect responding involved I AM-SELFISH-TRUE, etc.). In contrast, during the *inconsistent* trials, correct responding involved associating I AM with negative words and OTHERS with positive words (e.g. I AM-HONEST-FALSE; I AM-SELFISH-

TRUE; OTHERS ARE-SELFISH-FALSE; OTHERS ARE-HONEST-TRUE and incorrect responding involved I AM-SELFISH-FALSE, etc.).

*Phase 3: Post-experimental Measures.* Following completion of the IRAP, participants completed a second exposure to the four IRAP-related Response Scales. All participants completed the entire experiment in a single session that lasted between 20 and 30mins.

## RESULTS

*D<sub>IRAP-(trial-type)</sub> Data.* Consistent with the analyses conducted on the D scores in Experiment 3, the  $D_{IRAP-(trial-type)}$  scores were calculated for each participant in Experiment 4 on the two groups of IRAP trial-types that might be collectively referred to as Pro-Me (i.e. the I AM-Positive and I AM-Negative trial-types combined) and Pro-Others (i.e. the OTHERS ARE-Positive and OTHERS ARE-Negative trial-types combined). This data is presented in Figure 7. The data indicated that participants showed a strong implicit bias towards the self, with little or no opinion with regards to how they saw others.

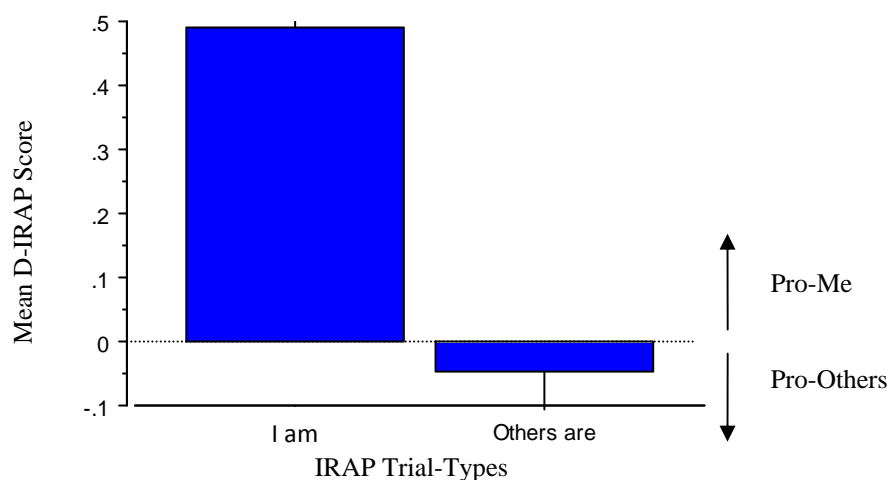


Figure 7. The mean  $D_{IRAP-(trial-type)}$  scores and standard error bars for the sample stimuli in Experiment 4.



A one-way repeated measures ANOVA revealed a large significant main effect for trial-type [ $F(1, 15) = 26.53, p = .0001, \eta_p^2 = .64$ ]. Two one sample t-tests indicated that the I AM trial-type produced a large significant difference from zero,  $t(31) = 6.39, p < .0001, \eta^2 = .73$ , but, as expected, OTHERS ARE was non-significant ( $p = .56$ ). Overall therefore, the IRAP indicated that participants had strong positive implicit self-regard, but had no strong opinions about others. Put simply, the positive self-regard did not imply that they had to confirm that others were positive or negative.

### *Explicit Measures*

*Psychological Measures.* The means scores for each of the DASS-21 subscales fell within the normal range: Depression ( $M = 8.0, SD = 8.67$ ; Norm  $< 9$ ); Anxiety ( $M = 6.63, SD = 3.85$ ; Norm  $< 7$ ); and Stress ( $M = 13.13, SD = 6.73$ ; Norm  $< 14$ ). As expected, participants also produced high RSES scores ( $M = 19.25; SD = 4.19$ ), indicating high levels of self-esteem, as well as high AAQ-II scores ( $M = 47.94, SD = 9.98$ ; Norm =  $50.72, SD = 9.19$ ), indicating high levels of psychological flexibility (high acceptance/low avoidance).

*Four IRAP-Related Response Scales.* Mean scores (and standard deviations) for Pre- and Post-IRAP ratings for the positive and negative IRAP words are presented in Table 10. As expected, participants overall rated the negative words as more negative, more uncomfortable and more unbelievable (i.e. less believable) and harder to experience than the positive words. Although there were some apparent changes on these scales from Pre- to Post-IRAP, these were small.

Table 10

*Mean Ratings (and Standard Deviations) for Groups of Positive and Negative Words at Pre- and Post-IRAP in Experiment 4*

IRAP-Related Response Scale	Pre-IRAP	Post-IRAP
<i>Feeling</i>		
Positive	5.69 (1.31)	5.60 (1.29)
Negative	2.39 (1.31)	2.32 (1.10)
<i>Discomfort</i>		
Positive	5.94 (1.28)	5.83 (1.32)
Negative	2.53 (1.33)	2.68 (1.31)
<i>Believability</i>		
Positive	3.18 (1.37)	3.31 (1.52)
Negative	4.71 (1.51)	4.44 (1.68)
<i>Willingness</i>		
Positive	5.42 (1.28)	5.55 (1.27)
Negative	2.96 (1.38)	3.14 (1.52)

*Note* .Higher scores indicate words were rated as more positive, more comfortable, less believable and participants were more willing to experience them.

Four separate 2x2 repeated measures ANOVAs were conducted (one for each of the IRAP-related Response Scales) with time (Pre-IRAP and Post-IRAP) and word type (positive and negative) as the within participant variables. The results from the scales were identical to the previous experiment. The *Feeling* Scale indicated a large significant main effect for word type [ $F(1, 30) = 126.64, p < .0001, \eta_p^2 = .81$ ], but not for time or the interaction ( $p$ 's = .65 and .97, respectively). This was also the case for the *Discomfort* Scale -- word type [ $F(1, 30) = 139.15, p < .0001, \eta_p^2 = .82$ ], time and interaction ( $p$ 's = .93 and .66, respectively); the *Believability* Scale -- word type [ $F(1, 30) = 10.83, p = .003, \eta_p^2 = .26$ ], time and interaction ( $p$ 's = .64 and .62, respectively); and the *Willingness* Scale -- word type [ $F(1, 30) = 76.75, p < .0001, \eta_p^2 = .72$ ], time and interaction ( $p$ 's = .54 and .94, respectively).

Similar to Experiment 3, participants in the current experiment rated the positive words as significantly more positive, more comfortable, more believable and easier to experience than the negative words. No differences were recorded between Pre- and Post-IRAP ratings on any of the four rating scales.

#### *Correlations between Implicit and Explicit Measures*

*Psychological Measures.* Correlational analyses were conducted between the three implicit measures (IRAP-Overall score, IRAP-I AM score and IRAP-OTHERS ARE score) and each of the psychological measures (depression, anxiety, stress, RSES and AAQ-II). The results indicated weak correlations in the majority of cases (all  $r$ 's < .23). However, a significant large positive correlation was obtained between the RSES scale and the IRAP overall score ( $r = .63, n = 16, p = .007$ ) and this effect was driven by responses to the I AM trial-types ( $r = .51, n = 16, p = .002$ ). Taken together, the results indicated that higher levels of explicit self-esteem in this context were associated with a stronger I AM-Positive implicit bias.

*Four IRAP-Related Response Scales.* Correlational analyses were conducted between the implicit measures (IRAP-Overall score, IRAP-I AM score and IRAP-OTHERS ARE score) and each of the four explicit IRAP-related Response Scales (Feeling, Discomfort, Believability and Willingness). The results indicated weak to moderate correlations in the majority of cases (all  $r$ 's range between .02 and .43).

## DISCUSSION

The results from Experiment 4 were very similar to those recorded in the previous study. On the IRAP, participants had a strong positive implicit self-regard,

but no strong opinions as to how they viewed others -- thus indicating that their positive self-regard was dependent on confirming that they were positive and not negative, but was not dependent on confirming that others were negative or positive. On the explicit measures, participants rated the positive words as more positive, more comfortable, more believable and easier to experience these than the negative words, and these ratings did not change from Pre- to Post-IRAP. Again, participants scored within the normal range on each of the psychological measures. No correlations emerged between the implicit measures and any of the four IRAP-related Response Scales. However, a large significant correlation did emerge between the implicit measure and the RSES, indicating that higher self-reported levels of self-esteem were associated with higher implicit levels of positive self-regard.

The results from the current IRAP indicated once again that participants had a strong positive implicit self-regard, although they were as ambivalent about others as they had been about what they themselves were not. This strong positive bias led us to generate one further IRAP in which such a positive self-bias might not emerge. This was the focus of Experiment 5.

## EXPERIMENT 5

### METHOD

#### *Participants*

Sixteen participants (all female) aged between 18 and 24 years ( $M = 21$ ) comprised the final sample of Experiment 5. All were undergraduate students, recruited through Faculty announcements at NUIM. All had a high level of fluency in English and had normal, or corrected-to-normal, vision. No incentives of any kind were offered for participation. In addition to the 16 participants who completed the experiment, four failed to reach the required criteria in either the training or test phases of the IRAP, hence their data was excluded from further analyses.

#### *Setting*

All aspects of the setting for Experiment 5 were identical to the previous IRAP experiments.

#### *Materials and Apparatus*

*Explicit Measures.* The three psychological measures and the four IRAP-related Response Scales used in the previous IRAPs were employed again in Experiment 5.

*IRAP.* The primary difference between the IRAP employed in previous experiments and Experiment 5 is that the current program presented I AM and I SHOULD BE as sample stimuli. The target stimuli and response options remained the same. Naturally, this also necessitated some changes to the instructions provided to

participants. The stimulus arrangements employed in Experiment 5 are presented in Table 11.

Table 11

*The Stimulus Arrangements Employed in Experiment 5*

<b>Sample Stimulus 1</b> I AM	<b>Sample Stimulus 2</b> I SHOULD BE
<b>Response Option consistent with Sample 1</b> TRUE	<b>Response Option consistent with Sample 2</b> FALSE
<b>Target Stimuli consistent with Sample 1</b>	<b>Target Stimuli consistent with Sample 2</b>
KIND CONFIDENT TRUSTING HONEST SECURE POPULAR	SELFISH SELF-CONSCIOUS JEALOUS FAKE INSECURE LONELY

*Procedure*

All procedural aspects and ethical guidelines of Experiment 5 were identical to the previous studies.

*Phase 1: Pre-Experimental Measures.* Participants first completed the DASS-21, the RSES and the AAQ-II, as well as the four IRAP-related Response Scales.

*Phase 2: The IRAP.* The IRAP for Experiment 5 consisted of four basic trial-types: I AM–Positive; I AM–Negative; I SHOULD BE–Positive; and I SHOULD BE –Negative (see Figure 8). Within each block of 24 trials, the sample stimulus I AM was presented on 12 trials, with I SHOULD BE presented on the other 12 trials. The presentation of sample stimuli was randomised throughout each block. In the presence of each sample, each of the twelve target words (six positive and six negative) appeared twice. The locations of TRUE and FALSE were randomised across trials.

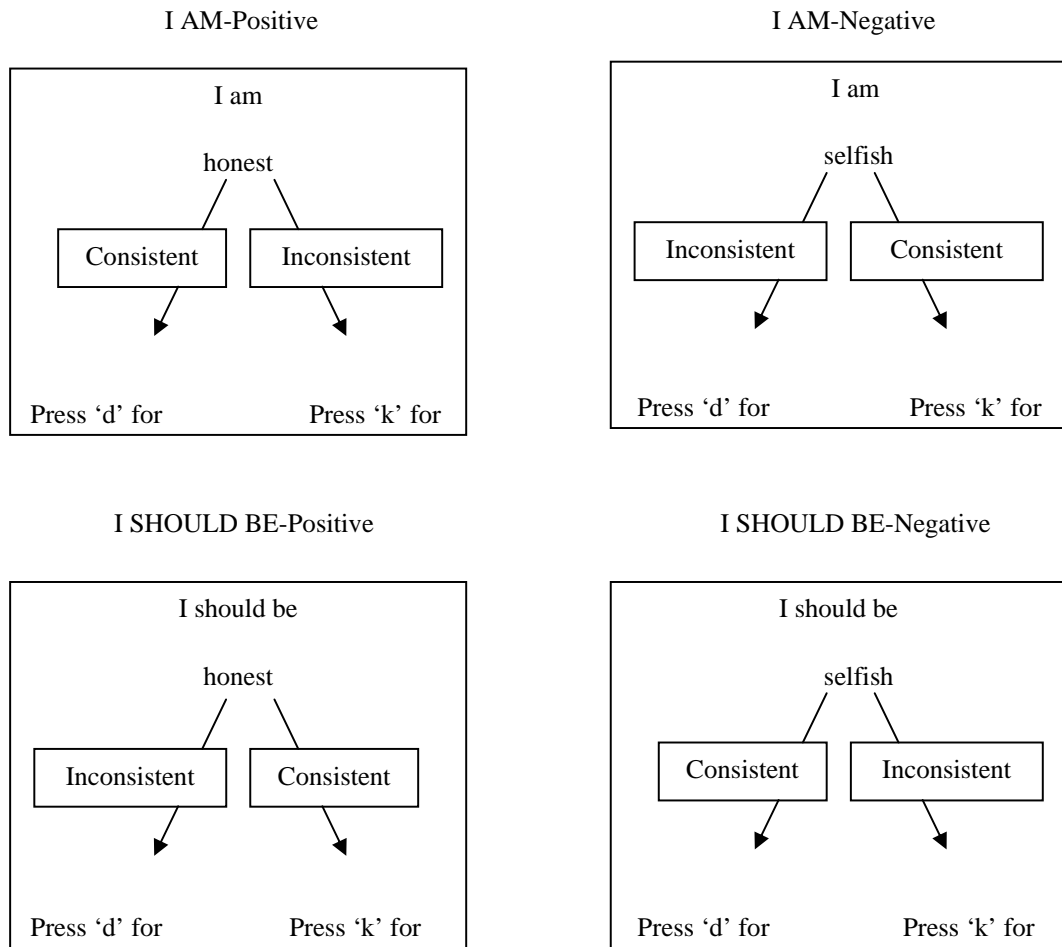


Figure 8. Examples of the four IRAP trial-types used in Experiment 5. The superimposed arrows with text boxes indicate the responses deemed consistent or inconsistent, but these boxes and arrows did not appear on screen during the experiment.

*Consistent versus Inconsistent Responding.* In the current IRAP, correct responses on *consistent* trials involved associating I AM with positive words and I SHOULD BE with negative words (e.g. I AM-HONEST-TRUE; I AM-SELFISH-FALSE; I SHOULD BE-SELFISH-TRUE; I SHOULD BE-HONEST-FALSE and incorrect responding involved I AM-HONEST-FALSE, etc.). In contrast, during the *inconsistent* trials, correct responding involved associating I AM with negative words and I SHOULD BE with positive words (e.g. I AM-HONEST-FALSE; I AM-

SELFISH-TRUE; I SHOULD BE -SELFISH-FALSE; I SHOULD BE -HONEST-TRUE and incorrect responding involved I AM-HONEST-TRUE, etc.).

*Phase 3: Post-experimental Measures.* Following completion of the IRAP, participants completed a second exposure to the four IRAP-related Response Scales. All participants completed the entire experiment in a single session that lasted between 20 and 30mins.

## RESULTS

*D<sub>IRAP-(trial-type)</sub> Data.* Consistent with the analyses conducted on the D scores in the previous IRAPs, the  $D_{IRAP-(trial-type)}$  scores were calculated for each participant in Experiment 5 on the two groups of IRAP trial-types referred to as Pro-Me (I AM-Positive and I AM-Negative) and Pro-Should be (I SHOULD BE-Positive and I SHOULD BE-Negative). This data is presented in Figure 9. Again, participants showed a strong implicit positive bias towards the self. They also indicated that they should not be positive (i.e. that they should be negative), but this effect was relatively weak.



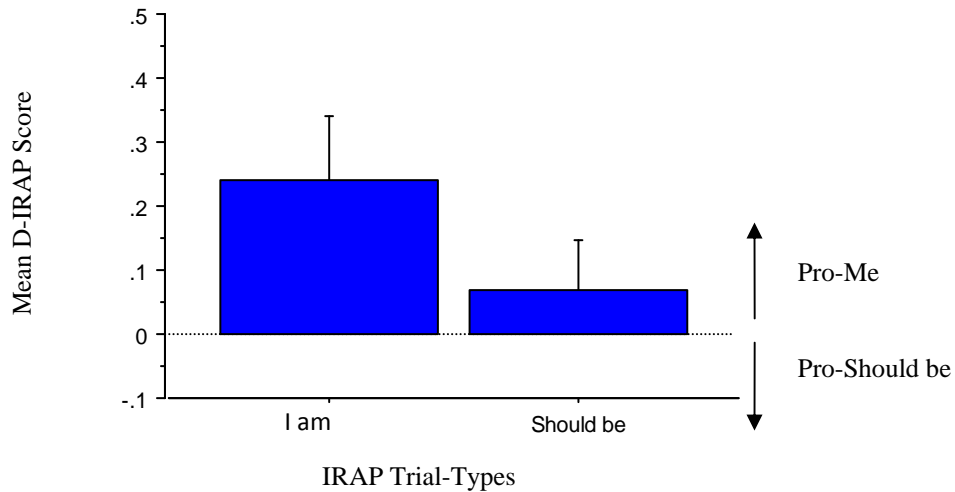


Figure 9. The mean  $D_{\text{IRAP-(trial-type)}}$  scores and standard error bars for the sample stimuli in Experiment 5.

A one-way repeated measures ANOVA revealed a significant main effect for trial-type [ $F(1, 15) = 7.50, p = .02, \eta_p^2 = .33$ ]. Two one sample t-tests indicated that the I AM trial-type produced a small significant difference from zero,  $t(31) = 2.39, p < .02, \eta^2 = .28$ , but I SHOULD BE was non-significant ( $p = .39$ ). Overall, the IRAP indicated that participants were certain that they were positive and were less strong in their conviction that they should be negative (i.e. I should not be positive). Put simply, the positive self-regard did not imply that they had to confirm that they should be positive or negative.

### Explicit Measures

*Psychological Measures.* The means scores for each of the DASS-21 subscales fell within the normal range: Depression ( $M = 4.88, SD = 4.13; \text{Norm} < 9$ ); Anxiety ( $M = 4.63, SD = 4.05; \text{Norm} < 7$ ); and Stress ( $M = 11.63, SD = 7.31; \text{Norm} < 14$ ). As expected, participants also produced high RSES scores ( $M = 21.06; SD = 4.14$ ), indicating high levels of self-esteem, as well as high AAQ-II scores ( $M =$

50.81,  $SD = 6.97$ ; Norm = 50.72,  $SD = 9.19$ ), indicating high levels of psychological flexibility (high acceptance/low avoidance).

*Four IRAP-Related Response Scales.* Mean scores (and standard deviations) for Pre- and Post-IRAP ratings for the positive and negative IRAP words are presented in Table 12. As expected, participants overall rated the negative words as more negative, more uncomfortable and more unbelievable (i.e. less believable) and easier to experience than the positive words. There were only minor differences Pre- to Post- IRAP.

Table 12

*Mean Ratings (and Standard Deviations) for Groups of Positive and Negative Words at Pre- and Post-IRAP in Experiment 5*

IRAP-Related Response Scale	Pre-IRAP	Post-IRAP
<i>Feeling</i>		
Positive	6.32 (.95)	6.38 (.95)
Negative	2.24 (1.13)	2.12 (.99)
<i>Discomfort</i>		
Positive	6.06 (1.04)	6.08 (1.01)
Negative	2.25 (1.30)	2.40 (1.16)
<i>Believability</i>		
Positive	2.39 (1.30)	2.37 (1.27)
Negative	5.48 (1.26)	5.70 (1.03)
<i>Willingness</i>		
Positive	6.26 (.79)	6.20 (1.02)
Negative	3.38 (1.70)	3.14 (1.79)

*Note.* Higher scores indicate words were rated as more positive, more comfortable, less believable and participants were more willing to experience them.

Four separate 2x2 repeated measures ANOVAs were conducted (one for each of the IRAP-related Likert Scales) with time (Pre-IRAP and Post-IRAP) and word type (positive and negative) as the within participant variables. The results were very similar to both previous experiments. The *Feeling* Scale indicated a large significant

main effect for word type [ $F(1, 30) = 595.59, p < .0001, \eta_p^2 = .95$ ], but not for time or the interaction ( $p$ 's = .79 and .61, respectively). This was also the case for: the *Discomfort Scale* -- word type [ $F(1, 30) = 120.15, p < .0001, \eta_p^2 = .80$ ], time and interaction ( $p$ 's = .93 and .49, respectively); the *Believability Scale* -- word type [ $F(1, 30) = 185.15, p < .0001, \eta_p^2 = .86$ ], time and interaction ( $p$ 's = .51 and .62, respectively); and the *Willingness Scale* -- word type [ $F(1, 30) = 118.71, p < .0001, \eta_p^2 = .80$ ], time and interaction ( $p$ 's = .66 and .75, respectively).

Similar to the previous IRAPs, participants in Experiment 5 again rated the positive words as significantly more positive, more comfortable, more believable and easier to experience than the negative words. No differences were recorded between Pre- and Post-IRAP ratings.

#### *Correlations between Implicit and Explicit Measures*

*Psychological Measures.* Correlation analyses were conducted between the three implicit measures (IRAP-Overall score, IRAP-I AM score and IRAP-I SHOULD BE score) and each of the psychological measures (depression, anxiety, stress, RSES and AAQ-II). Overall, the results indicated a pattern of medium to large correlations between the IRAP Overall score and each of the five psychological measures. Specifically, large negative correlations were obtained between the IRAP overall score and reported levels of stress ( $r = -.62, n = 16, p = .008$ ) and between the IRAP overall score and reported levels of depression ( $r = -.54, n = 16, p = .03$ ). These results indicate that the higher the IRAP overall score (i.e. higher implicit positive self regard), the lower the levels of stress and depression. All remaining  $r$ 's < .46).

To further investigate the cause of the effects for the stress and depression scales, the correlations obtained between both psychological measures and the two

groups of trial-types (IRAP-I AM and IRAP-I SHOULD BE) were examined.

However, all correlations in this regard were weak (all  $r$ 's < -.28).

*Four IRAP-Related Response Scales.* Correlational analyses were conducted between the implicit (IRAP-Overall score, IRAP-I AM score and IRAP-I SHOULD BE score) and each of the four explicit IRAP-related Response Scales (Feeling, Discomfort, Believability and Willingness). The results indicate weak correlations in all cases (all  $r$ 's < .29).

## DISCUSSION

The results from Experiment 5 indicated that participants had a strong positive implicit self-regard, but had no strong opinions on how they felt they should be -- indicating that their positive self-regard was dependent on confirming how they saw themselves as being, but was not dependent on how they felt they should be. The results of the explicit measures indicated that participants overall scored within the normal range on each of the psychological measures. As expected with the explicit measures, participants also rated the positive words as more positive, more comfortable, more believable and easier to experience than the negative words. The ratings did not change from Pre- to Post-IRAP. No correlations emerged between the implicit measures and any of the four IRAP-related Response Scales. However, large significant negative correlations did emerge between the implicit measure and self-reported levels of stress and depression, indicating that higher levels of positive implicit self-regard was associated with significantly lower levels of stress and depression.

## EXPERIMENTS 3, 4 AND 5

### COMPARATIVE ANALYSES

*Implicit Measures.* Across each of the three experiments, participants reported a strong implicit positive self-regard, as indicated by responses on the Pro-Me trial-types. However, the strength of the effect for Pro-Me was different across the three experiments. The results of a one-way between groups ANOVA indicated that participants presented with the OTHERS ARE stimuli (Experiment 4) produced the largest effect size for the I AM trial-types and this was significantly larger than the effect for the I SHOULD BE stimuli ( $p = .05$ ), which produced the smallest effect size (Experiment 5). However, those presented with the I AM NOT stimuli (Experiment 3) did not significantly differ from the other two ( $p$ 's  $> .16$ ).

*Psychological Measures.* A series of one-way between groups ANOVAs (with experiment as the between groups variable) was conducted for each of the psychological measures. The results indicated no significant differences between the participants across the three experiments on any of the measures (all  $p$ 's  $> .90$ ). Thus, the differences in the Pro-Me responses across all three experiments could not be attributed to differences on the explicit psychological measures.

Despite the groups providing a similar pattern of self-reported levels of depression, anxiety, stress, self-esteem and acceptance/avoidance, differences did emerge across the three experiments in terms of correlations between the implicit ratings and the psychological measures. Specifically, the implicit results recorded in Experiment 3 (I AM versus I AM NOT) produced no significant correlations with any of the psychological measures. In contrast, the results from Experiment 4 (I AM versus OTHERS ARE) indicated that higher levels of positive implicit self-regard

were associated with higher levels of self-esteem. Different again was the pattern of correlations reported in Experiment 5 (I AM versus I SHOULD BE), with higher levels of positive implicit self-regard associated with lower levels of stress and depression.

*IRAP-Related Response Scales.* Across all three experiments, participants rated the positive words as significantly more positive, more comfortable, more believable and easier to experience than the negative words. Across all three experiments, no changes were recorded between Pre- to Post-IRAP ratings. Furthermore, no correlations were reported between the implicit measure and the four scales in any of the three experiments.

## DISCUSSION

The primary aim of Experiments 3, 4 and 5 was to create an IRAP that would tap into implicit self-relevant content in a manner that would subsequently allow us to manipulate these implicit attitudes through a brief defusion intervention. The results from Experiment 5 (I AM versus I SHOULD BE) produced the significantly weakest implicit positive self-regard, relative to the other two IRAPs. In addition, implicit responses produced in this context also correlated with explicit levels of stress and depression. On this basis, it was considered that the I SHOULD BE statements offered the best potential implicit measure for use with a defusion intervention.

Across all three experiments, positive and negative words were utilised as the target stimuli. However, it is arguable that the use of more complete positive and negative *statements* may better reflect the types of thoughts experienced by clients within a therapeutic setting, and therefore would better represent the target of a

defusion intervention. Thus, the next phase of the current research program was twofold. (1) It was necessary to design an IRAP (using the I SHOULD BE stimuli) that incorporated positive and negative statements as opposed to simple positive and negative words. (2) It was necessary to determine whether or not the implicit measures were sensitive to a defusion manipulation. To this end, it was deemed necessary to compare IRAP performances for groups of individuals with more or less experience with defusion (i.e. a sample of students versus a sample of ACT therapists).

## **Chapter 5**

### **Experiments 6 and 7**

**Using the IRAP to Measure Reactions**

**to Self-Statements:**

**Comparing Undergraduate Students**

**with Therapists**



## **Chapter 5**

### **Experiments 6 and 7**

#### **Using the IRAP to Measure Reactions to Self-Statements: Comparing Undergraduate Students with Therapists**

##### INTRODUCTION

The results from the three IRAP experiments in Chapter 4 indicated that undergraduate participants showed strong implicit positive self-regard on all three types of self-IRAP presentation. Although these results were entirely as one would expect, we were somewhat surprised by the lack of ‘doubt’ in their responses overall. Although the presentation in Experiment 5 (I AM versus I SHOULD BE) was potentially the best of the three formats for subsequent use in the context of defusion, it remained the case that even here participants had little doubt about how positively they perceived themselves to be. One possible reason for these limited effects overall may concern the use of single words as target stimuli, which may not sufficiently tap into the complexities of one’s sense of self. As an analogy, one might consider that differences in reaction times are often recorded for IAT studies that use pictures rather than words as stimuli (e.g. Dasgupta, McGhee, Greenwald, & Banaji, 2000). Hence, the presentation of self-statements on-screen might yield different IRAP effects from single words, especially as one thinks about oneself. For example, individuals may respond differently to “lonely” compared with “I am so alone that it hurts”.

One of our primary aims in the studies thus far was to develop an IRAP that might be potentially sensitive to a defusion intervention and it was anticipated that the

results from the use of self-statements (rather than words) in Experiment 6 would bring us one step closer towards achieving this aim. However, the results from our earlier work in delivering defusion interventions clearly highlighted the complexities involved in doing this in an experimental context. Hence in Experiment 7, we turned our attention towards a group of participants who already had a long history of defusion. Specifically, prior to attempting to manipulate the IRAP outcomes with defusion interventions directly, we sampled a group of ACT therapists, whom we assumed had substantive histories of defusion with regard to clients, as well as their own psychological content. Although it was difficult in this context to predict precisely what type of outcomes ACT therapists would produce on a self-IRAP, we predicted that it would differ to some extent from undergraduates who had no extended history of defusion.

### *The Current Research*

The current chapter presents two separate experiments. Experiment 6 was designed to assess the impact of using positive and negative self-statements, rather than words, as target stimuli in the IRAP. In order to generate the target stimuli, the six positive and six negative words used previously were developed into fuller statements. To systematically compare the impact of words versus statements, Experiment 6 involved two groups of undergraduates -- one exposed to a Word-IRAP, the other exposed to a Statement-IRAP.

Experiment 7 was designed to investigate the potential sensitivity of the IRAP to a defusion intervention. The primary manipulation here was the differential level of experience with defusion between undergraduate students and ACT therapists, when both were presented with self-statements in an IRAP.

## EXPERIMENT 6

### METHOD

#### *Participants*

A total of 32 female participants (all female) aged between 18 and 28 ( $M = 22$ ) comprised the final sample of Experiment 6. Half of these participants ( $n = 16$ ) were presented with a Word-IRAP that was identical to the presentation in Experiment 5. The other half were presented with a Statement-IRAP developed from the words used in the previous study. All were undergraduate students, recruited through Faculty announcements at NUIM. All participants had a high level of fluency in English and normal, or corrected-to-normal, vision. There were no incentives for participation. In addition to the 32 participants who completed the experiment, nine participants (four of which were assigned to the Word-IRAP, with the remaining assigned to the Statement-IRAP) failed to reach the required criteria in either the training or test phases of the IRAP, hence their data was excluded from further analyses.

#### *Setting*

All aspects of the setting for Experiment 6 were identical to the previous IRAPs.

#### *Materials and Apparatus*

*Explicit Measures.* All participants completed the DASS-21 (Lovibond, & Lovibond, 1995 -- see Appendix G and Chapter 3) and the RSES (Rosenberg, 1965 -- see Appendix I and Chapter 4), used in the previous studies.

*IRAP*. Two separate *IRAP* procedures were employed in Experiment 6, namely a Word-*IRAP* and a Statement-*IRAP*. The primary difference between these related to the target stimuli. The stimuli employed in the Word-*IRAP* were identical to Experiment 5 (see Table 11, Chapter 4). The stimuli employed in the Statement-*IRAP* are presented in Table 13. Naturally, the change in target stimuli also necessitated some changes to the instructions provided to participants.

Table 13

*The Stimulus Arrangements Employed in the Statement-IRAP in Experiment 6*

<b>Sample Stimulus 1</b> I AM	<b>Sample Stimulus 2</b> I SHOULD BE
<b>Response Option consistent with Sample 1</b> TRUE	<b>Response Option consistent with Sample 2</b> FALSE
<b>Target Stimuli consistent with Sample 1</b>	<b>Target Stimuli consistent with Sample 2</b>
Warm-hearted in nature	Unpleasant to be around
Happy-go-lucky	Afraid to show who I really am
Open to new relationships	Jealous of those around me
Comfortable with who I am	Lying to myself and others
Confident in my abilities	Constantly seeking reassurance
Surrounded by people I love	So alone that it hurts

*Procedure*

All ethical guidelines for Experiment 6 were identical to the previous *IRAP*s. A number of minor modifications were made to the procedural aspects of the current experiment. Specifically, due to the different nature of the stimuli employed (i.e. words versus statements), participants were not presented with the *IRAP*-related Response Scales. As such, there were only two experimental phases.

*Phase 1: Pre-Experimental Measures.* Participants first completed the DASS-21 and the RSES.

*Phase 2: The IRAP.* The Word-IRAP for Experiment 6 consisted of four basic trial-types. These were identical to those outlined for Experiment 5. The Statement-IRAP also consisted of four basic trial-types: I AM–Positive; I AM–Negative; I SHOULD BE–Positive; and I SHOULD BE –Negative (see Figure 10). Within each block of 24 trials, the sample stimulus I AM was presented on 12 trials, with I SHOULD BE presented on the other 12 trials. The presentation of sample stimuli was randomised throughout each block. In the presence of each sample, each of the twelve target *statements* (six positive and six negative) appeared twice. The locations of TRUE and FALSE were randomised across trials.

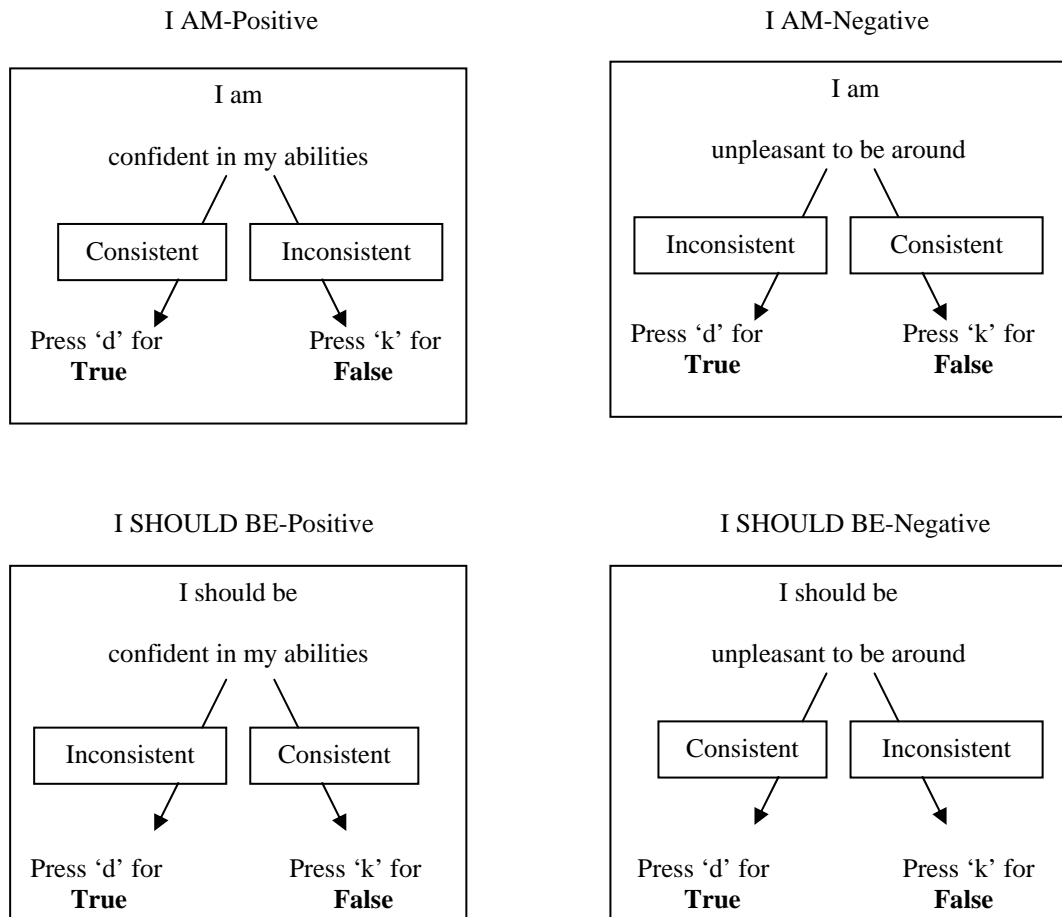


Figure 10. Examples of the four Statement-IRAP trial-types used in Experiment 6. The superimposed arrows with text boxes indicate the responses deemed consistent or inconsistent, but these boxes and arrows did not appear on screen during the experiment.

*Consistent versus Inconsistent Responding.* For the Word-IRAP, correct responses on *consistent* trials involved associating I AM with positive words and I SHOULD BE with negative words and *inconsistent* trials involved associating I AM with negative words and I SHOULD BE with positive words. This pattern of responding is identical to Experiment 5.

For the Statement-IRAP, correct responding on *consistent* trials involved associating I AM with positive statements and I SHOULD BE with negative statements (e.g. I AM-CONFIDENT IN MY ABILITIES-TRUE; I AM-

UNPLEASANT TO BE AROUND-FALSE; I SHOULD BE-UNPLEASANT TO BE AROUND-TRUE; I SHOULD BE-CONFIDENT IN MY ABILITIES-FALSE and incorrect responding involved I AM-CONFIDENT IN MY ABILITIES-FALSE, etc.). In contrast, during the *inconsistent* trials, correct responding involved associating I AM with negative statements and I SHOULD BE with positive statements (e.g. I AM-CONFIDENT IN MY ABILITIES-FALSE; I AM-UNPLEASANT TO BE AROUND-TRUE; I SHOULD BE-UNPLEASANT TO BE AROUND-FALSE; I SHOULD BE-CONFIDENT IN MY ABILITIES-TRUE and incorrect responding involved I AM-CONFIDENT IN MY ABILITIES-TRUE, etc.).

All participants completed the current experiment in a single session that lasted between 20 and 30mins.

## RESULTS

*D<sub>IRAP-(trial-type)</sub> Data.* Consistent with the analyses conducted on the D scores in the previous IRAPs, the  $D_{IRAP-(trial-type)}$  scores were calculated for each participant in both the Word-IRAP and Statement-IRAP groups in Experiment 6. In both cases, the IRAP trial-types were referred to as Pro-Me (I AM-Positive and I AM-Negative) and Pro-Should be (I SHOULD BE-Positive and I SHOULD BE-Negative). This data is presented in Figure 11. Participants showed a strong implicit positive bias towards the self, with little or no opinion with regards to how they felt they should be. This was true for both Word- and Statement- IRAP groups.

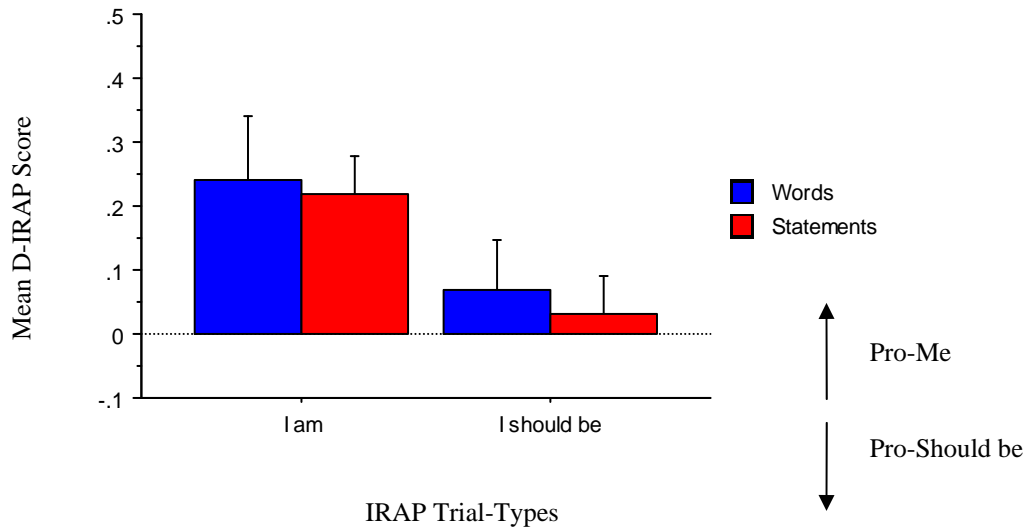


Figure 11. The mean  $D_{\text{IRAP-(trial-type)}}$  scores and standard error bars for the Word-IRAP and Statement-IRAP in Experiment 6.

A 2x2 mixed-between-within AVOVA was conducted, with group (Word-IRAP and Statement-IRAP) as the between participants variable and trial-type (I AM and I SHOULD BE) as the within participants variable. The results revealed a significant main effect for trial-type [ $F(1, 15) = 8.32, p = .007, \eta_p^2 = .20$ ], but not group and the interaction effect was also non-significant ( $p$ 's = .60 and .91, respectively).

Four separate one sample t-tests indicated that the I AM trial-type produced a large significant difference from zero for both the Word-IRAP,  $t(31) = 2.39, p = .02, \eta^2 = .41$  and Statement-IRAP,  $t(39) = 3.64, p = .0008, \eta^2 = .28$ . As expected, the effect for the I SHOULD BE trial-type was non-significant for both the Word-IRAP ( $p = .39$ ) and the Statement-IRAP ( $p = .61$ ).

Overall therefore, the results from both IRAPs indicated that participants had strong positive implicit self-regard. Participants were slightly less certain of this in the context of the Statement-IRAP than the Word-IRAP, however, these differences were non-significant. Participants had no strong opinions with regard to what they thought



they should be. Although participants were slightly more certain that they should be negative in the context of the Word-IRAP stimuli, the differences between groups were non-significant.

*Explicit Measures*

*Psychological Measures.* The means scores (and standard deviations) for each of the psychological measures for both the Word-IRAP and Statement-IRAP groups are presented in Table 14. Scores from both groups fell within the normal range for the depression, anxiety and stress sub-scales. As expected, both groups reported high levels of self-esteem.

Table 14

*Means (and Standard Deviations) for Each of the Psychological Measures for the Word-IRAP and Statement-IRAP Groups in Experiment 6*

Psychological Measure	Norms	Word-IRAP	Statement-IRAP
RSES	-	19.25 (4.19)	22.60 (4.68)
<i>DASS</i>			
Depression	< 9	8.0 (8.67)	4.50 (6.08)
Anxiety	< 7	6.63 (3.85)	3.70 (6.33)
Stress	< 14	13.13 (6.73)	9.60 (5.99)

*Note.* ‘-’ indicates that no normative data is available.

A series of independent samples t-tests were conducted for each of the psychological measures (with IRAP group as the independent samples grouping variable). The results revealed no significant difference between scores for the Word-IRAP group and the Statement-IRAP group for any of the psychological measures (all  $p$ 's > .31). Thus, the differences obtained in the D-scores for the I AM and I

SHOULD be trial-types and in the pattern of correlation between implicit and explicit measures could not be accounted for by differences in initial self-report ratings.

### *Correlations between Implicit and Explicit Measures*

Correlation analyses were conducted between the three implicit measures (IRAP-Overall score, IRAP-I AM score and IRAP-I SHOULD BE score) and each of the psychological measures (depression, anxiety, stress and RSES). The results from the Word-IRAP indicated large negative correlations between the IRAP overall score and levels of stress ( $r = -.62, n = 16, p = .008$ ), and between the IRAP overall score and levels of depression ( $r = -.54, n = 16, p = .03$ ). These results indicate that the higher the IRAP overall score (i.e. higher implicit positive self regard), the lower the levels of stress and depression (all remaining  $r$ 's  $< .46$ ). Separate correlational analyses were then conducted between the psychological measures and the two groups of trial-types (IRAP-I AM and IRAP-I SHOULD BE). However, all correlations were weak (all  $r$ 's  $< -.28$ ). For the Statement-IRAP, only small to medium correlations were obtained between any of the implicit and psychological measures (all  $r$ 's  $< .32$ ).

## DISCUSSION

The results from Experiment 6 indicated that both words and statements revealed strong positive implicit self-regard from participants. However, subtle differences emerged between the two groups in terms of the strength of their reactions. Specifically, participants presented with the Statement-IRAP were slightly less certain that they were positive and were slightly less certain that they should be negative than

the reactions observed on the Word-IRAP. These differences could not be accounted for by differences in the explicit psychological measures.

The pattern of responding across the two groups suggests that the Statement-IRAP created more uncertainty in terms of how certain participants were that they were indeed positive, as well as being less certain that they should not be positive. It was therefore considered to be the IRAP that would be most appropriate for use with a defusion intervention. Furthermore, it seemed likely that the whole statements better reflected the types of psychological content with which individuals struggle and as such was a better procedure to employ as a clinical analogue.

The next phase of the current research program was designed to determine if the Statement-IRAP would indeed be susceptible to changes as a result of defusion. An initial step in this process was to determine whether or not there would be any difference in IRAP performance between groups of individuals with more or less experience with defusion. This was the primary focus of Experiment 7. Specifically, we investigated performances on the Statement-IRAP of a group of undergraduate participants, compared to a group of ACT therapists. The former group were considered to have no experience of defusion, while the latter group were considered to have considerable experience with defusion, either personally or through therapeutic work with clients.

## EXPERIMENT 7

### METHOD

#### *Participants*

*Non-ACT Group:* Twenty participants (all female) aged between 22 and 29 ( $M = 24$ ) comprised the final sample for the Non-ACT Group. All were undergraduate students, recruited through Faculty announcements at NUIM. All participants had a high level of fluency in English and normal, or corrected-to-normal, vision. There were no incentives for participation. In addition to the 20 participants that completed the study, five failed to reach the required criteria in either the training or test phases of the IRAP, hence their data was excluded from further analyses.

*ACT Group.* Eighteen participants (12 female) aged between 24 and 51 ( $M = 33$ ) comprised the final sample for the ACT Group. All were ACT therapists/trainee therapists, recruited through on an online advertisement placed on the official ACT email forum ([acceptanceandcommitmenttherapy@yahoogroups.com](mailto:acceptanceandcommitmenttherapy@yahoogroups.com)). All participants had a high level of fluency in English and normal, or corrected-to-normal, vision. There were no incentives for participation. A total of 52 ACT members volunteered to participate. Of these, 28 did not attempt the IRAP experiment and did not contact the experimenter again. Of the participants that did attempt the experiment, six failed to reach the required criteria in either the training or testing phase of the IRAP.

#### *Setting*

*Non-ACT Group.* Participants completed the experiment in one of the Experimental Cubicles in the Psychology Department at NUIM. Only one participant was present in a cubicle at any time. All individuals completed all aspects of the

experiment individually. The experimenter remained seated outside the cubicle at all times and only interacted with the participants during the first IRAP practice blocks.

*ACT Group.* Participants completed all aspects of the experiment electronically, with details passed between experimenter and participant via email (see Procedure for specific details). Participants were asked to complete the experiment in a single setting, in a quiet room that was free from any distraction.

### *Materials and Apparatus*

*Psychological Measures.* All participants in both the ACT and Non-ACT Groups completed three psychological measures, namely the DASS-21 (Lovibond, & Lovibond, 1995 -- see Appendix G and Chapter 3), the RSES (Rosenberg, 1965 -- see Appendix I and Chapter 4) and the Philadelphia Mindfulness Scales (PHLMS; Cardaciotto, Herbert, Forman, Moitra, & Farrow, in press -- Appendix L).

The *PHLMS* is a 20-item inventory that provides a measure of current levels of mindfulness. It consists of two sub-scales (awareness and acceptance), each of which contains 10 questions. Individuals are asked to rate how often they had experienced each statement over the past week on a scale of 1 (never) to 5 (very often). The awareness sub-scale is comprised of all the odd numbered questions, none of which are negatively reversed. The acceptance sub-scale is comprised of all the even numbered items, all of which are negatively reversed. Total scores for each sub-scale range from 10 to 50, with higher scores on each representing higher levels of awareness and higher levels of acceptance, respectively. Mean scores obtained with a non-clinical undergraduate sample are as follows: awareness:  $M = 36.65$ ,  $SD = 4.93$  and acceptance:  $M = 30.19$ ,  $SD = 5.84$ ). Both scales indicate good psychometric

properties, with good reliability and internal consistency (Cronbach's alpha for awareness and acceptance sub-scales = .81 and .85, respectively).

*IRAP-Related Believability Scale.* Participants in both groups completed an IRAP-related Believability Scale (see Appendix M). This scale was designed to assess participants' explicit attitudes to the self-relevant stimuli that would be presented during the IRAP. The scale measured participants' levels of believability associated with the 12 self-statements employed in the IRAP-- six positive statements and six negative statements. Response options ranged from -3 (extremely unbelievable) to +3 (extremely believable).

*IRAP.* The IRAP employed in Experiment 7 was identical to the Statement-IRAP from Experiment 6 (see Table 13). Although the IRAP for the ACT Group was presented through an online website, all aspects of the IRAP program were identical to those presented to the undergraduate participants who completed the experiment at NUIM.

### *Procedure*

All procedural aspects and ethical guidelines of Experiment 7 were identical to previous experiments.

*Non-ACT Group.* Participants in this group completed three experimental phases, namely: (1) *Pre-Experimental Measures* (brief demographic questionnaire, psychological measures and IRAP-related Believability Scales); (2) the *IRAP*; and (3) *Post-Experimental Measures* (IRAP-related Believability Scales).

*ACT Group.* Participants in the ACT Group completed the three experimental phases as outlined above. However, it was necessary to provide participants with additional information and instructions in order to successfully execute an online version of the experiment. Specifically, once participants had volunteered to take part

(by responding to the advertisement placed on the ACT email forum), they were provided with an experimental guide containing necessary information to complete the experiment (see Appendix N). Contained in the experimental guide were a personalised identification number and password that were used to access a website, used previously for other types of experimental research studies (<https://rft.ori.org/exper/webPresentation.php?opt=4>).

This guide was emailed to participants and the specific details for each of the four steps outlined in the guide were provided as four separate documents attached to the email. On completion of all four steps, participants were required to return the signed consent form as well as responses to the Pre- and Post-Experimental measures. The IRAP data was automatically recorded on the website and made available for collection immediately after each participant had completed the IRAP.

*Pre-Experimental Measures.* Participants were required to electronically complete a brief demographic questionnaire to elicit information regarding: age; gender; previous experience with the IRAP and experience in ACT by simply typing their responses in spaces provided. In addition, participants completed the psychological measures and IRAP-related Believability Scales. In order to complete these scales electronically, a red X had been placed at the end of each questionnaire item. Participants were provided with the following instructions:

In order to respond to each item on the rating scale, you must drag the and drop the **X** over the number that most accurately reflects your response to that particular item

To drag and drop the **X**

- Double click on the **X** relevant to the question you are responding to.
- A shaded square should appear around the **X**.
- Move the cursor so that it hovers over one of the *sides* of the square (and *not* the corner of the square).
- Hold down the left mouse button and drag the mouse to the number that represents the response you want to make.
- Release the left mouse button.
- The **X** should now appear over the response that you wish to make.

Having completed these ratings, participants were then instructed to proceed immediately to the IRAP.

*IRAP.* Participants in the ACT Group completed an identical IRAP to that outlined for the Statement-IRAP in Experiment 6. The IRAP procedure was accessed through an online web link provided at the end of the instruction manual (<https://rft.ori.org/exper/webPresentation.php?opt=4>). At this point, participants were required to enter their unique personalised identification number and password in order to proceed with the IRAP task. At the end of the IRAP, a message appeared indicating that the sorting tasks were complete and they were free to close down their internet connection. Only data from the test phase was saved on the internet site and was accessible online to the experimenter, who also had a unique identification number and password to access the website.

*Post-Experimental Measures.* Participants were required to complete a second exposure to the IRAP-related Believability Scales in an identical manner as outlined above.

Participants were then asked to save the changes they had made to the consent form, as well as the Pre- and Post-Experimental measures and to return these to the experimenter via email.

Upon receipt of these documents, the experimenter replied to each participant via email with a full debrief of the experimental procedure. Participants were also provided with the opportunity to ask any further questions at this time. No participants had any further questions, however, a number requested a copy of the overall results once all the data had been collected and written for publication.

All participants completed the study in a single experimental session. Participants in the Non-ACT Group completed all aspects of the experiment between



20 and 30mins. in total. Participants in the ACT Group generally took longer to complete the experiment, averaging 45 to 55mins. to complete.

## RESULTS

*D<sub>IRAP-(trial-type)</sub> Data.* Consistent with the analyses conducted on the D scores in the previous experiments, the  $D_{IRAP-(trial-type)}$  scores were calculated for each participant in Experiment 7 on the two groups of IRAP trial-types that might be collectively referred to as Pro-Me and Pro-Should be for both the ACT Group and the Non-ACT Group. This data is presented in Figure 12. Both groups of participants showed a strong implicit positive bias towards the self, with little or no opinion on how they felt they should be. Interestingly, participants in the ACT Group appeared more certain that they were positive than those in the Non-ACT Group.

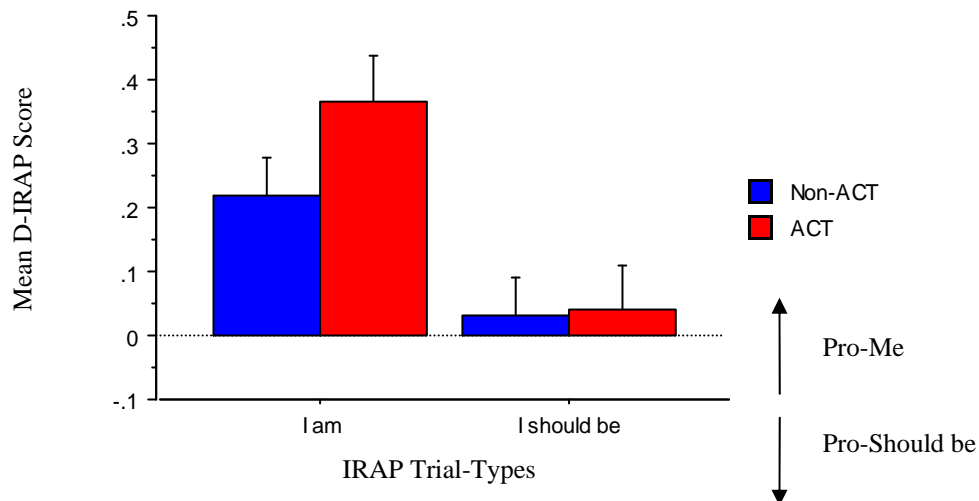


Figure 12. The mean  $D_{IRAP-(trial-type)}$  scores and standard error bars for the sample stimuli for the Non-ACT and ACT groups in Experiment 7.

A 2x2 mixed-between-within AVOVA was conducted, with group (ACT and Non-ACT) as the between participants variables and trial-type (I AM and I SHOULD BE) as the within participants variable. The results revealed a significant main effect for trial-type [ $F(1, 36) = 13.94, p = .007, \eta_p^2 = .28$ ]. However, the main effect for group and the interaction effect were both non-significant ( $p$ 's = .28 and .32, respectively).

Four separate one sample t-tests indicated that the I AM trial-type produced a large significant difference from zero for both the ACT Group,  $t(35) = 4.938, p < .0001, \eta^2 = .59$  and Non-ACT Group,  $t(39) = 3.64, p = .0008, \eta^2 = .47$ . As expected, the effect for the I SHOULD BE trial-type was non-significant for both the ACT Group ( $p = .54$ ) and the Non-ACT Group ( $p = .61$ ).

Overall therefore, the results indicated that both groups had a strong positive implicit self-regard, although participants in the ACT Group were slightly more certain of this. Both groups reacted similarly to the I SHOULD BE statements, with no strong opinions being shown in this regard.

### *Explicit Measures*

*Psychological Measures.* The means scores (and standard deviations) for each of the psychological measures are presented in Table 15. Scores from both groups fell within the normal range for the depression, anxiety and stress sub-scales. As expected, both groups reported high levels of self-esteem, as measured by the RSES. Both groups also reported high levels of awareness and acceptance on the mindfulness measure.

Table 15

*Means (and Standard Deviations) for Each of the Psychological Measures for the ACT and Non-ACT Groups in Experiment 7*

Psychological Measure	Norms	ACT Group	Non-ACT Group
RSES	-	25.29 (4.06)	22.60 (4.68)
<i>DASS</i>			
Depression	< 9	4.59 (4.17)	4.50 (6.08)
Anxiety	< 7	1.29 (1.57)	3.70 (6.33)
Stress	<14	9.65 (5.67)	9.60 (5.97)
<i>PHLMS</i>			
Awareness	36.65 (4.93)	39.53 (7.20)	32.50 (5.01)
Acceptance	30.19 (5.84)	40.94 (4.59)	32.75 (4.27)

Note. '-' indicates that no normative data is available.

A series of separate unpaired samples t-tests were conducted to determine whether there was any difference between the two groups in terms of their performance on the psychological measures. Results indicated a significant difference between the two groups in terms of levels of awareness,  $t(35) = 3.49$ ,  $p = .001$ ,  $\eta^2 = .25$ , and acceptance,  $t(35) = 5.62$ ,  $p < .0001$ ,  $\eta^2 = .47$ . As expected, participants in the ACT Group produced higher scores on both measures. Differences between the two groups on the remaining psychological measures were non-significant (all  $p$ 's  $> .07$ ).

*IRAP-Related Believability Scale.* Mean scores (and standard deviations) for responses to the IRAP-related Believability Scale for both the ACT and Non-ACT groups are presented in Table 16. As expected, participants in both groups rated the positive statements as more believable than the negative statements. Although there were some apparent changes on these scales from Pre- to Post-IRAP for both groups, these differences were small.

Table 16

*Mean Scores (and Standard Deviations) for Believability Ratings of Positive and Negative Statements at Pre- and Post-IRAP for ACT and Non-ACT Groups in Experiment 7*

Believability Scale	Pre-IRAP		Post-IRAP	
	ACT	Non-ACT	ACT	Non-ACT
Positive	1.80 (1.28)	1.82 (1.02)	1.78 (1.66)	2.03 (.755)
Negative	-1.32 (1.76)	-1.38 (1.61)	-1.50 (1.85)	-1.37 (1.72)

*Note.* Higher scores indicate statements were rated as more believable.

A 2x2x2 mixed-between-within repeated measures ANOVA was conducted, with group (ACT versus Non-ACT) as the between participant variable and statement type (positive and negative) and time (Pre-IRAP and Post-IRAP) as the within participant variables. The results indicated a large significant main effect for statement type [ $F(3, 102) = 86.11, p < .0001, \eta_p^2 = .72$ ]. All other effects were non-significant (all  $p$ 's  $> .59$ ).

Thus, participants rated the positive statements as significantly more believable than the negative words. However, no differences were recorded between Pre- and Post-IRAP ratings and this effect held for both the ACT and Non-ACT Groups.

#### *Correlations between Implicit and Explicit Measures*

*Psychological Measures.* Separate correlation analyses were conducted between the three implicit measures (IRAP-Overall score, IRAP-I AM score and IRAP-I SHOULD BE score) and each of the psychological measures (depression, anxiety, stress, RSES, awareness and acceptance) for both groups. Results from the ACT Group revealed small to medium correlations between the psychological

measures and the IRAP scores (all  $r$ 's  $< -.32$ ). Correlational analyses from the Non-ACT Group also revealed small to medium correlations (all  $r$ 's  $< -.40$ ).

*IRAP-Related Believability Scale.* Separate correlation analyses were conducted between the three implicit measures (IRAP-Overall score, IRAP-I AM score and IRAP-I SHOULD BE score) and the believability IRAP-related Believability Scale for both groups. Results from the ACT Group revealed small correlations between the believability ratings and the IRAP scores (all  $r$ 's  $< .22$ ). Correlational analyses from the Non-ACT Group also revealed small correlations (all  $r$ 's  $< .08$ ).

## DISCUSSION

The results from Experiment 7 indicated that participants in both the ACT and Non-ACT Groups had a strong positive implicit self-regard. However, subtle differences emerged between the two groups in terms of the strength of their reactions. Specifically, participants in the ACT Group were more certain that they were positive than the Non-ACT Group. Participants in both groups indicated no strong reaction towards the I SHOULD BE statements. Interestingly, the ACT Group also had significantly higher levels of awareness and acceptance on the PHLMS. However, neither measure correlated with any of the implicit measures. Correlations between the implicit measures and the explicit believability ratings of the positive and negative statements were also non-significant.

## GENERAL DISCUSSION

Experiment 6 was designed to investigate the relative impact of words versus statements in a sample of undergraduate students. Overall, the statements appeared to be associated with greater uncertainty in terms of positive implicit self-regard, as well as in terms of how participants felt they should be. It was therefore considered that the statements would be more beneficial than words for use with a defusion intervention.

Experiment 7 was designed as a preliminary investigation into the impact of defusion on an individual's implicit self-regard. Specifically, participants with a history of experience with defusion (ACT Group) were compared to those with no such experience (Non-ACT Group). The results indicated subtle differences between the two groups in terms of implicit positive self-regard and explicit levels of awareness and acceptance, with the ACT group scoring higher in both contexts. Overall, the pattern of responding in Experiment 7 suggested that those participants with a greater history of defusion were more certain that they were positive and not negative at an implicit level. However, there are a number of other potential variables that could influence this pattern of responding. For example, although participants in the ACT Group had a greater history of defusion techniques, they were also familiar with a range of other ACT techniques and were professionals (as opposed to undergraduates), thus, making it difficult to decipher the specific contribution of defusion history to these results.

The next and final phase of the current research program was designed to investigate the impact of a simple defusion intervention on implicit ratings of the self in a sample of individuals with no prior experience of defusion. This would also afford an investigation into the specific contribution of defusion to the different

pattern of responding obtained for the ACT and Non-ACT Groups in Experiment 7.

This was the primary focus of Experiment 8, the final experiment of the current thesis.

## **Chapter 6**

### **Experiment 8**

#### **Using the IRAP to Measure Reactions to Positive and Negative Self-Statements Following a Cognitive Defusion Intervention**



## **Chapter 6**

### **Experiment 8**

# **Using the IRAP to Measure Reactions to Positive and Negative Self-Statements Following a Cognitive Defusion Intervention**

## INTRODUCTION

Experiment 7 provided a preliminary investigation into the impact of defusion on implicit reactions to positive and negative self-statements. The results of this experiment revealed a different pattern of responding in individuals with a greater history of defusion than those with no history of defusion, with the former reporting stronger implicit positive self-regard. With this in mind, the current experiment was designed to investigate the impact of defusion on implicit self-regard in a sample of individuals with no history of defusion, where there remained some potential room for improvement in implicit self-esteem.

### *Implicit Cognitions and Clinical Interventions*

Implicit measures have a limited history of use in measuring the impact of clinical interventions and little evidence exists of their utility in this regard. Nevertheless, research investigating the impact of clinical interventions on implicit cognitions is warranted, particularly given the argument that various psychopathologies may comprise of dysfunctional implicit beliefs that may not be consciously available and yet may still influence clinically relevant behaviours (Beck,

Rush, Shaw, & Emery, 1979; de Houwer, 2002). For example, research has identified the role of implicit cognitions in a number of clinically relevant areas including: anxiety-related behaviours (Egloff, & Schmuckle, 2002); suicidal behaviour (Nock, & Banaji, 2007) and substance abuse (e.g. cocaine: Wiers, Houben, & de Kraker, 2007; marijuana: Ames et al., 2007; and alcohol: Thush, & Wiers, 2007). In many cases, implicit measures have predicted these behaviours over and above explicit measures. As such, implicit measures provide an important method of assessing clinically relevant thoughts that may otherwise remain untreated (de Jong et al., 2001; Robinson, Meier, Zetocha, & McCaul, 2005).

As a result, researchers have begun to investigate the impact of clinical interventions on implicit cognitions. One of the first investigations into the utility of implicit procedures as a measure of the impact of clinical intervention was reported by Teachman and Woody (2003). Specifically, the authors investigated the impact of a group-based exposure treatment on implicit spider associations in a sample of spider phobic individuals. The results from four separate IATs indicated significant decreases in associations between spiders/disgusting and spiders/afraid from Pre- to Post-Intervention and these effects remained at a two-month follow-up. However, the associations between spiders/bad and spiders/dangerous did not change following the intervention. Importantly, no changes in IAT scores were recorded for the No Intervention control group at either post-test or follow-up.

More recently, researchers have begun to investigate the impact of a specific pain intervention on implicit self-pain associations as well as on implicit self-esteem (Grumm, Erbe, von Collani, & Nestler, 2008). Specifically, patients with chronic pain were exposed to an intensive pain treatment every day over a four week period. The results indicated a decrease in implicit associations between pain and the self for the

chronic pain group, whereas no such decrease was reported in the group of healthy controls (none of which were exposed to any intervention). Interestingly, participants in the pain group reported significantly lower levels of implicit self-esteem at Pre-Intervention relative to the control group. Post-Intervention scores indicated no change in implicit self-esteem for the pain group. Unexpectedly, a significant decrease in implicit self-esteem was reported for the control group. Similar positive results have also been reported in the context of social anxiety, with reductions reported on an anxiety-related IAT in a sample of socially anxious individuals at Post-intervention ratings (Gamer, Schmukle, Luka-Krausgrill, & Egloff, in press).

### *The Current Study*

Experiment 8 was focused on measuring the specific impact of cognitive defusion on implicit responses to a range of positive and negative self-statements and as such, was designed to determine whether implicit measures may offer a supplemented measure of defusion to explicit self-reports. The primary manipulation in this experiment was exposure to a defusion intervention or exposure to a placebo condition. Specifically, IRAP performance was measured before and after exposure to one of the conditions in a sample of university students (undergraduates and postgraduates). Once again, as in the previous IRAPs, Experiment 8 was purely exploratory, and as such we made no firm predictions regarding the outcomes.

## METHOD

### *Participants*

Twenty-four participants (10 males) aged between 22 and 30 years ( $M = 26$ ) comprised the final sample of Experiment 8. Half of these participants ( $n = 12$ ) were randomly assigned to the Defusion Group, with the other half assigned to the Placebo Group (5 males and 7 females in each group). Participants were either undergraduates or postgraduate students, recruited through Faculty announcements at NUIM. All participants had a high level of fluency in English and normal, or corrected-to-normal, vision. There were no incentives for participation. In addition to the 24 participants who completed the experiment, seven failed to reach the required criteria in either the training or test phases of the IRAP. Hence, their data was excluded from further analyses.

### *Setting*

The study was conducted in the experimental cubicles in the Psychology Department at NUIM. Noise and distraction were kept to a minimum throughout the experiment. Each participant completed all aspects of the experiment individually and in an individual cubicle. The experimenter remained seated outside the cubicle for the duration of the study.

### *Materials and Apparatus*

*Explicit Measures.* Participants completed the DASS-21 (Lovibond, & Lovibond, 1995 -- see Appendix G and Chapter 3), the RSES (Rosenberg, 1965 -- see Appendix I and Chapter 4) and the PHLMS (Cardaciotto et al., in press -- see

Appendix L and Chapter 5), as well as the IRAP-related Believability Scale employed in Experiment 7 (see Appendix M and Chapter 5).

*IRAP.* Two identical IRAPs were used in the current experiment, one at Pre-Intervention and one at Post-Intervention. Both IRAPs were identical to the Statement-IRAP outlined in Experiments 6 and 7 (see Chapter 5).

*Therapeutic Scripts.* The therapeutic scripts for the Defusion and Placebo Conditions employed here were identical to those used in Experiment 2. Specifically, the Defusion Group here was identical to the Defusion/Defusion Condition in Experiment 2. Participants in the Defusion Group, therefore, were exposed to both a defusion rationale and defusion exercise. The Placebo Group was identical to the Placebo/Placebo Condition in Experiment 2. Participants assigned to this group were exposed to both a placebo rationale and a placebo exercise.

In addition, participants were also presented with a stopwatch and tape recorder, which were essential to the completion of these experimental exercises.

### *Ethical Issues*

Experiment 8 followed the same ethical guidelines as the previous IRAP experiments and the earlier experiments involving defusion interventions.

### *Procedure*

Experiment 8 contained ten experimental phases, always completed in the same order. The ten phases are summarised in Table 17.

Table 17

*An Outline of the Ten Experimental Phases employed in Experiment 8*

Experimental Phase	Description
Phase 1	Pre-Experimental psychological measures (DASS-21, RSES, and PHLMS) and IRAP-related Believability Scale
Phase 2	Pre-Intervention Statement-IRAP
Phase 3	Pre-Intervention ratings on the IRAP-related Believability Scale
Phase 4	Generating a personally relevant uncomfortable thought and restating this thought as a single word
Phase 5	Pre-Intervention ratings of personally relevant thought for levels of discomfort, believability and willingness
Phase 6	Participants were exposed to either the Defusion rationale and exercise or the Placebo rationale and exercise
Phase 7	Post-Intervention ratings of personally relevant thought for levels of discomfort, believability and willingness
Phase 8	Post-Intervention adherence ratings assessing honesty in responding and level of difficulty associated with instructions
Phase 9	Post-Intervention Statement-IRAP
Phase 10	Post-Intervention ratings on the IRAP-related Believability Scale

Participants were thanked for their co-operation and time. They were then appropriately debriefed about the nature of the experiment and invited to ask any questions. Although made available to them, no participants opted for short breaks at any point. All participants completed the study in a single session that lasted between 75 and 90mins.

## RESULTS

*D<sub>IRAP-(trial-type)</sub> Data.* Consistent with the analyses conducted on the D scores in previous experiments, the  $D_{IRAP-(trial-type)}$  scores were calculated for each participant in Experiment 8 for the Pro-ME and the Pro-Should be trial-types for both the Defusion Group and the Placebo Group at both Pre- and Post-Intervention intervals. This data is presented in Figure 13. For the I AM trials, both groups produced identical Pre-Intervention ratings indicating a strong positive implicit self-regard. However, the groups produced different Post-Intervention scores for the I AM trial-types. Although both groups became more certain that they were positive, those assigned to the Defusion Group produced a larger effect. Although both groups indicated that they should be positive, Pre-Intervention scores in this regard were very small. Post-Intervention scores for this trial-type suggested that those assigned to the Placebo Group became more certain that they should be positive than those assigned to the Defusion Group.

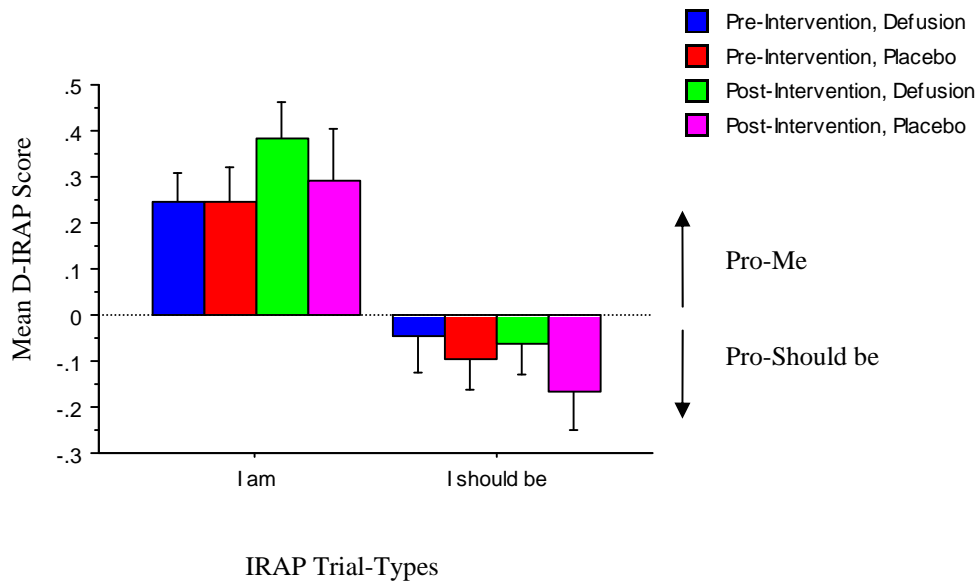


Figure 13. The mean  $D_{\text{IRAP-(trial-type)}}$  scores and standard error bars for the pre- and post-intervention IRAPs for the Defusion and Placebo groups in Experiment 8.

A 2x2x2 mixed-between-within repeated measures ANOVA was conducted, with group (Defusion and Placebo) as the between participants variable and trial-type (I AM and I SHOULD BE) and time (Pre- and Post-Intervention) as the within participant variables. The results revealed a significant main effect for trial-type [ $F(1, 44) = 38.63, p < .0001, \eta^2 = .48$ ]. However, the main effect for group and time were both non-significant ( $p$ 's = .26 and .65, respectively), as were the interaction effects (all  $p$ 's < .23).

A series of separate one sample t-tests indicated that the I AM trial-type produced a large significant for both groups at both Pre- and Post-Intervention (Defusion Group: Pre-Intervention =  $t(23) = 4.03, p = .0005, \eta^2 = .60$ ; Post-Intervention =  $t(23) = 4.79, p < .0001, \eta^2 = .68$  and Placebo Group: Pre-Intervention =  $t(23) = 3.22, p = .004, \eta^2 = .49$ ; Post-Intervention =  $t(23) = .264, p = .02, \eta^2 = .58$ ). As expected, all effects for the I SHOULD BE trial-type were non-significant (all  $p$ 's > .06).



Overall therefore, the results indicate that both groups had a strong positive implicit self-regard prior to, and following, exposure to the intervention. Both groups generally reported no strong opinions towards the I SHOULD BE statements, although Post-Intervention ratings for the Placebo group approached significance -- this effect was not obtained for the Defusion Group.

### *Explicit Measures*

*Psychological Measures.* The mean scores (and standard deviations) for each of the psychological measures are presented in Table 18. Scores from both groups fell within the normal to moderate range for the depression and stress sub-scales. Furthermore, participants in the Defusion Group reported moderate levels of anxiety, while those in the Placebo Groups reported normal levels. Despite these differences, participants in both groups reported high levels of self-esteem, as measured by the RSES. Both groups also reported high levels of awareness and acceptance on the mindfulness measure.

Table 18

*Mean Scores (and Standard Deviations) for Each of the Psychological Measures for the Defusion and Placebo Conditions in Experiment 8*

Psychological Measure	Norms	Defusion	Placebo
RSES	-	21.50 (3.40)	22.33 (5.98)
<i>DASS</i>			
Depression	< 9	7.67 (6.76)	6.6 (5.73)
Anxiety	< 7	11.33 (6.63)	5.83 (7.46)
Stress	<14	13.83 (8.29)	9.83 (6.95)
<i>PHLMS</i>			
Awareness	36.65 (4.93)	36.50 (4.62)	34.25 (5.99)
Acceptance	30.19 (5.84)	29.60 (6.84)	36.25 (6.18)

*Note.* ‘-’ indicates that no normative data is available.

A series of separate independent samples t-test were conducted to determine whether there was any difference between the two groups in terms of their performance on the psychological measures. Results indicated a significant difference between the two groups in terms of levels of acceptance,  $t(22) = -2.54, p = .02, \eta^2 = .23$ , with participants in the Defusion Group producing significantly lower scores on the acceptance scales. Differences between the two groups on the remaining psychological measures were non-significant (all  $p$ 's > .07).

*IRAP-Related Believability Scale.* Mean scores (and standard deviations) for responses to the IRAP-related Believability Scale for both the Defusion and Placebo Groups across the three time periods are presented in Table 19.

Table 19

*Mean Scores (and Standard Deviations) for Believability Ratings of Positive and Negative Statements at Times 1, 2 and 3 for Defusion and Placebo Groups in Experiment 8*

Believability Scale	Time 1 (Pre-Experimental)		Time 2 (Pre-Intervention)		Time 3 (Post-Intervention)	
	Defusion	Placebo	Defusion	Placebo	Defusion	Placebo
Positive	1.68 (1.17)	1.85 (1.40)	1.64 (1.36)	2.18 (.91)	1.81 (1.11)	2.12 (1.21)
Negative	-1.04 (1.61)	-1.46 (1.76)	-1.03 (1.58)	-1.57 (1.67)	-1.28 (1.70)	-1.68 (1.53)

*Note.* Higher scores indicate statements were rated as more believable.

A 2x2x3 mixed-between-within repeated measures AVOVA was conducted, with group (Defusion and Placebo) as the between participant variable and statement type (positive and negative) and time (Pre-Experimental, Pre-Intervention and Post-Intervention) as the within participant variables. The results indicated a large significant main effect for statement type [ $F(1, 63) = 144.91, p < .0001, \eta_p^2 = .70$ ].

All other effects were non-significant (all  $p$ 's > .29). Thus, participants rated the positive statements as significantly more believable than the negative statements. However, believability ratings did not change across the three time periods and this effect held for both the Defusion and Placebo Groups.

### *Correlations between Implicit and Explicit Measures*

*Psychological Measures.* Separate correlation analyses were conducted between the three implicit measures (IRAP-Overall scores, IRAP-I AM score and IRAP-I SHOULD BE score) and each of the psychological measures (depression, anxiety, stress, RSES, awareness and acceptance) for both groups. Results from the Defusion group at the Pre-Intervention IRAP indicated small to medium correlations (all  $r$ 's < .32). Interestingly, Post-Intervention correlations revealed a number large correlations emerged between the Post-Intervention IRAP-Overall score and some of the psychological measures. Specifically, the Post-Intervention IRAP-Overall score produced large negative correlations with levels of anxiety ( $r = -.82, n = 12, p = .0006$ ) and depression ( $r = -.76, n = 12, p = .003$ ), as well as a large positive correlation with levels of self-esteem ( $r = .68, n = 12, p = .01$ ).

Results from the Placebo Group at Pre-Intervention IRAP overall indicated small to medium correlations (all  $r$ 's < .44), with the exception of the large negative correlation obtained between the IRAP-Overall score and levels of acceptance ( $r = -.57, n = 12, p = .05$ ). Interestingly, this correlation did not hold at Post-Intervention ( $r = .009$ ). All remaining correlations between Post-Intervention IRAP scores and psychological measures were small to medium in size (all  $r$ 's < .47).

Taken together, therefore, the results indicated a different pattern of correlation for both groups at Pre- and Post-Intervention ratings, suggesting a change

in pattern of implicit responding following exposure to either Defusion or Placebo instructions.

*IRAP-Related Believability Scale.* Separate correlation analyses were conducted between the three Pre-Intervention implicit measures (IRAP-Overall scores, IRAP-I AM score and IRAP-I SHOULD BE score) and the IRAP-related Believability Scale for both groups. Results from the Defusion Group revealed a large negative correlation between the believability ratings of the positive statements and implicit responses to the I SHOULD BE trial-types ( $r = -.57, n = 12, p = .003$ ). That is, participants with high explicit rating of believability of positive were less likely to implicitly indicate that they should be positive. All remaining correlations were small to medium (all  $r$ 's  $< .41$ ). Correlational analyses from the Placebo Group revealed only small to medium correlations (all  $r$ 's  $< .41$ ).

*Explicit Ratings of Self-Generated Thought.* Consistent with the analysis in Experiment 2, a series of analyses were conducted on self-reported levels of discomfort, believability and willingness associated with the self-generated uncomfortable thought at Pre- and Post-Intervention.

*Discomfort.* The mean discomfort ratings at Pre- and Post-Intervention intervals for the Defusion and Placebo Groups are presented in Table 20. Participants' Pre-Intervention scores ranged between 65 and 75 on the scale, where 100 represented maximum discomfort. In both cases, discomfort decreased from Pre- to Post-Intervention.

A 2x2 mixed-between-within repeated measures ANOVA was conducted on the discomfort data, with group (Defusion and Placebo) as the between participant variable and time (Pre- and Post-Intervention) as the within participant variable.

Results indicated a significant main effect for time [ $F(1, 22) = 19.67, p = .0002, \eta_p^2$

= .47], but not for group ( $p = .72$ ). The interaction effect was also non-significant ( $p = .33$ ). Mean scores from both Defusion and Placebo indicated that self-report levels of discomfort associated with the negative self-referential thought decreased from Pre- to Post-Intervention.

Table 20

*Means (and Standard Deviations) for Discomfort, Believability and Willingness Ratings at Pre- and Post-Intervention in Experiment 8*

Condition	Pre-Intervention	Post-Intervention
<i>Discomfort Ratings</i>		
Defusion	72.50 (17.12)	47.50 (22.61)
Placebo	65.00 (24.68)	49.17 (25.75)
<i>Believability Ratings</i>		
Defusion	66.67 (23.48)	50.00 (31.04)
Placebo	72.50 (23.01)	45.83 (25.75)
<i>Willingness Ratings</i>		
Defusion	39.17 (21.57)	45.83 (27.78)
Placebo	45.00 (30.30)	60.00 (22.16)

*Believability.* The mean believability ratings suggested an overall decrease in levels of believability from Pre- to Post-Intervention (see Table 20). Participants' Pre-Intervention scores ranged between 65 and 75 on the scale, where 100 represented maximum believability. The 2x2 ANOVA revealed a significant main effect for time [ $F(1, 22) = 12.31, p = .002, \eta_p^2 = .57$ ], with group and interaction effects non-significant ( $p = .92$  and  $.43$ , respectively). Mean scores indicated that self-reported levels of believability decreased from Pre- to Post-Intervention for both groups.

*Willingness.* The mean willingness ratings suggested an overall increase in levels of willingness from Pre- to Post-Intervention (see Table 20). Participants' Pre-Intervention scores ranged between 35 and 45 on the scale, where 100 represented

maximum willingness. The 2x2 ANOVA revealed a significant main effect for time [ $F(1, 22) = 6.48, p = .02, \eta_p^2 = .23$ ], with group and interaction effects non-significant ( $p = .31$  and  $.34$ , respectively). Mean scores indicated that self-reported levels of willingness increased from Pre- to Post-Intervention for both groups.

### *Adherence Measures*

*Honesty.* Results indicated that both groups responded very honestly when rating their self-generated negative thought (Defusion:  $M = 90.83, SD = 10.84$  and Placebo:  $M = 93.33, SD = 6.51$ , where 100 represents maximum honesty). Unsurprisingly, the results of an independent t-test confirmed that the two groups did not differ significantly from each other in this regard ( $p = .50$ ).

*Difficulty Ratings.* Overall, participants did not find the Defusion or Placebo instructions difficult to follow (Defusion:  $M = 29.17, SD = 24.66$  and Placebo:  $M = 32.50, SD = 33.61$ , where 100 represents maximum difficulty). Once again, the results of an independent t-test revealed no significant difference between the two groups in this regard ( $p = .78$ ).

## DISCUSSION

Experiment 8 investigated the impact of a simple defusion intervention on implicit ratings of the self in a sample of individuals with no prior experience of defusion. The results indicated that participants in both the Defusion and Placebo Groups had a strong implicit positive self-regard prior to exposure to the intervention. However, subtle differences emerged between the two groups at Post-intervention with participants in the Defusion Group becoming slightly more certain that they were

positive than the Placebo Group. Participants in both groups indicated no strong reactions towards the I SHOULD BE statements at Pre-Intervention. However, once again, subtle differences emerged between the two groups at Post-intervention. Specifically, participants in the Placebo Group became more certain that they should be positive than those assigned to the Defusion Group.

The results obtained here for the Defusion Group were consistent with those obtained in Experiment 7 in terms of an increased positive implicit self-regard associated with defusion. Despite this, subtle differences emerged elsewhere between the results obtained in the current study and those reported previously. First differences emerged between the results obtained here and those obtained in Experiment 2 in terms of explicit ratings of the self-generated personally relevant thought. Specifically, participants in the Defusion Group here reported increased willingness to experience their negative thought, however, no changes in willingness were recorded in Experiment 2. Furthermore, participants in the Placebo Group reported significantly decreased explicit discomfort and believability and increased willingness, however, no such changes were obtained for the Placebo/Placebo condition in Experiment 2.

Second, a large negative correlation emerged between the believability ratings of the positive statements and the I SHOULD BE implicit scores produced by the Defusion Group. Similar correlations were not previously reported. Third, a different pattern of correlations emerged between the implicit and explicit psychological measures at Pre- and Post-Intervention. Specifically, the Defusion Group did not produce any correlations at Pre-Intervention. However, Post-Intervention produced large negative correlations with anxiety and depression and large positive correlations with explicit self-esteem. In contrast, the Placebo Group produced large negative

correlations between the implicit measure and acceptance at Pre-Intervention. However, the Post-Intervention IRAP for this group was not correlated with any psychological measures. This pattern of correlations was again different to that obtained previously.

Experiment 8 marks the end of the experimental work for the current thesis. The next chapter (Chapter 7) presents a summary of the work conducted, as well as a general discussion of the findings of all studies in the current thesis. Chapter 7 also presents suggestions for future research emanating from the current work.



## **Chapter 7**

### **General Discussion**

## Chapter 7

### General Discussion

Research suggests that acceptance, rather than avoidance, of troublesome psychological content is associated with positive psychological health and well-being. Yet there remains limited knowledge on the processes through which acceptance-based strategies work. In the current thesis, we have argued that cognitive defusion is one core strategy that facilitates psychological acceptance of thoughts, feelings and emotions. The primary focus of the present experimental work, therefore, was to explore the conditions under which defusion works best and to investigate the underlying processes associated with it.

#### *Summary of Findings from Part I: Chapters 2 and 3*

Part I of the current thesis was designed to investigate the impact of cognitive defusion on different types of psychological content, as well as to determine the conditions that best facilitate this process. Specifically, the experimental work presented in Part I of the thesis investigated the self-reported emotional impact of cognitive defusion across a range of different experimental conditions.

Experiment 1 (Chapter 2) had three main aims: (1) to investigate the impact of placing the defusion prefix “I am having the thought that” in front of a series of self-statements, and as such offer an insight into the impact of an alternative defusion strategy than had previously been employed by Masuda et al. (2004); (2) to investigate the impact of simple defusion-related instructions; and (3) to investigate the relative impact of defusion on positive and negative self-statements.

The results of Experiment 1 indicated that: (1) The defusion prefix “I am having the thought that” significantly reduced levels of discomfort associated with the negative self-statements. This finding was consistent with the word repetition defusion strategy employed by Masuda et al. (2004). Unexpectedly, Experiment 1 also indicated that defusion *increased* the believability of the negative self-statements. This was inconsistent with Masuda et al. who reported that defusion *decreased* believability. Experiment 1 also reported that defusion increased willingness to experience the negative self-statements. This was the first empirical evidence of such an effect. (2) The results suggested the greater utility of a defusion exercise over defusion-related instructions, with the latter showing little or no impact on the emotional ratings of the negative self-statement. Again, this is consistent with the Masuda study. (3) The results indicated differences between the impact of defusion on positive relative to negative self-statements. Specifically, defusion had little impact on the former relative to the latter. Once again, this was the first manipulation of its kind.

Experiment 2 was designed to account for some of the discrepancies that arose between Experiment 1 and the previously published study by Masuda et al. (2004) (e.g. the believability ratings). It was also designed to account for possible shortcomings that arose from the design of Experiment 1. For example, unlike Masuda, we had not employed personally relevant stimuli. In addition, both the instructions and intervention employed here were brief. In response to these issues, Experiment 2 had three main aims: (1) to investigate the impact of defusion on personally generated negative self-relevant thoughts, as opposed to generic self-statements; (2) to investigate the impact of more extensive defusion-related instruction and exercise (i.e. word repetition) as opposed to the simple instructions

and brief exercise used previously; and (3) to investigate the relative impact of Defusion versus Thought Control and Placebo conditions.

The results from Experiment 2 indicated that: (1) similar to the previous research, defusion decreased discomfort. In addition, defusion in this context also decreased believability associated with the negative self-referent. Interestingly, defusion here had no impact on levels of willingness to experience negative thoughts. (2) Once again, the results overall suggested the slight superiority of experiential exercise over instructions, although instructions alone did impact ratings in certain conditions. (3) Participants exposed to some element of defusion (i.e. either defusion rationale and/or exercise) reported the largest decreases in discomfort and believability relative to both Thought Control and Placebo conditions.

### *Theoretical Issues from Part I: Chapters 2 and 3*

#### *The Emotional Impact of Defusion*

*Discomfort.* Both Experiments 1 and 2 indicated that defusion was associated with a significant decrease in the level of *discomfort* associated with negative thoughts about the self. This is consistent with the results obtained in the previous published research by Masuda et al. (2004), who also reported a decrease in discomfort ratings. Taken together, the results suggest that irrespective of which defusion exercise is employed, experientially engaging with defusion reduces discomfort of negative thoughts. From an ACT perspective, therapists do not attempt to alter the discomfort generated by thoughts and feelings and as such, we made no initial predictions regarding the impact of defusion on discomfort. Hence, we were initially surprised that defusion decreased discomfort. Nevertheless, consider now a context that also involved decreased believability and increased willingness.

Specifically, when one's perspective changes from focusing on the outcome of language to focusing on the process of language (i.e. from a perspective in which one *is* their thoughts to one in which they *have* their thoughts), then it is perhaps unsurprising that discomfort begins to dissipate.

*Believability.* Discrepancies emerged across studies in terms of the impact of defusion on levels of *believability* associated with the thoughts. Specifically, both Experiment 2 and the Masuda et al. (2004) study reported that defusion *decreased* levels of believability associated with the negative self-referents. In contrast, Experiment 1 reported that defusion *increased* believability. Perhaps, this was simply because personally-relevant content was not targeted in Experiment 1, but was in the other two studies. Or, it may be because different defusion techniques were employed. Alternatively, one might even argue that a defusion effect did not occur in Experiment 1. However, consider first that participants were required to respond to the discomfort and willingness scales by responding to the *whole* defusion statement (e.g. how uncomfortable does the statement "I am having the thought that I am a bad person" make you feel?). It seems likely, therefore, that participants responded to the believability of the statements in the same way (e.g. how believable is it that you are having the thought that you are a bad person?). In effect, the increased believability ratings for the defused statements indicated that participants believed that they were indeed having that thought, rather than the thought being true. In this sense, therefore, the increased believability ratings could be seen as evidence *for* a defusion effect as specified by ACT, especially when taken in the context of decreased discomfort and increased willingness for the same set of statements. Of course, we can only speculate about whether participants here were responding to the self-statement per se or to the fully defused statement.

*Willingness.* Discrepancies also emerged in relation to the impact of defusion on willingness to experience the negative self-referents. Specifically, Experiment 1 reported an *increase* in levels of willingness, where as in Experiment 2 willingness remained unaffected (Masuda et al. did not include a measure of willingness in their study). Again, it is possible that the difference emerged as a result of the different defusion exercises used across both studies. Or again, it is equally plausible that the differences emerged due to the type of psychological content targeted. Specifically, in Experiment 1, participants were presented with generic negative statements about the self, while in Experiment 2 they generated their own. So there are three generic possibilities. First in Experiment 1, participants were forced to experience painful statements and the defusion prefix helped with this. As a result, willingness went from low to high. The possibility that had the participants not been forced to use these statements and could instead choose others (as in Experiment 2), explains why there was an increase in willingness in Experiment 1 but not in Experiment 2. Second, it is simply possible that the fact that participants in Experiment 2 started out with relatively high willingness and thus outcomes on this measure could not really be expected to increase (i.e. a type of ceiling effect). Third, the content targeted in Experiment 1 was not personal, so it might be argued that it was easy for the defusion technique to make this content easier to bear. In contrast, the content of Experiment 2 was hard to bear and thus it makes sense that perhaps participants in the latter case would not get more willing to bear this.

#### *Instructions versus Exercise*

*Defusion.* The experimental work in Part I of the current thesis adds to the debate regarding the most effective delivery of defusion techniques (i.e. instructions

versus exercises). The results from Experiment 1 suggested the superiority of a defusion exercise and indicated that instructions had no impact on discomfort, believability or willingness for negative self-statements. These results were consistent with those obtained by Masuda et al. (2004) who reported no impact with a defusion rationale alone. Nevertheless, there were a number of experimental shortcomings associated with each of these attempts to assess the relative impact of instructions versus exercise. Specifically, in Experiment 1, the defusion-related instructions were very brief. Furthermore, adherence to these instructions was low. In the rationale alone condition in the Masuda et al. study, participants were instructed to apply the rationale to their negative thought but at no point did they practice the exercise with their target content. As a result, it is difficult to determine what participants actually did following the defusion instructions – for example some participants may have been able to successfully apply the strategy to their thoughts, whereas others may not. Experiment 2 of the current thesis was designed to account for these shortcomings by systematically comparing instructions versus exercises, whilst using a more complete set of defusion-related instructions. In this context, defusion decreased discomfort for both instructions and exercise when presented together or when either was presented with a placebo component. With believability, the defusion exercise was superior to instructions.

*Thought Control.* Experiment 2 also compared the impact of instructions versus exercises in the context of Thought Control. The results indicated that when presented alone (i.e. with Placebo), neither component impacted on discomfort, believability or willingness ratings. Once again, these results were different from those obtained by Masuda et al. (2004), who reported that thought control instructions decreased believability of negative self-referents. In part, this discrepancy may have

arisen from differences between the two studies. Specifically, Masuda et al. presented participants with a number of different control strategies (e.g. breathing, positive imagery and positive self-talk) compared to the single strategy (i.e. word replacement) presented in Experiment 2.

Taken together, the data suggest that defusion and thought control operate in different ways from each other in terms of the relative efficacy of instructions versus exercises. Furthermore, the impact of instructions and exercises depends on the outcome measurement. That is, instructions and exercises have different impacts in different areas. One is better for discomfort, the other for believability. In any case, it is almost an impossible question to answer because it is nearly impossible to have an exercise without some levels of instructions. Nevertheless, it seems to be the case that instructions alone without an exercise are likely less effective than with one.

#### *The Psychological Content with which Defusion Works Best*

One of the main aims of the current thesis was to determine the conditions under which defusion works best. To this extent, the experimental data presented in Part I contributes to this knowledge. The results from Experiment 1 suggested that defusion works best with negative psychological content, but has little impact on positive content. However, it might be argued that if individuals are fully defused from their psychological content, then they would be defused from both their negative *and* their positive thoughts. For example, if a client was ‘fully defused’ s/he would not only *not* believe negative thoughts, but would also not need to believe positive thoughts. On the contrary, one might equally argue that if defusion is viewed as a prophylactic against negative or troublesome psychological content, then it is not surprising that it should have little impact on positive content, because the latter does



not pose the same level of difficulty. The data here lend general support to the latter assumption.

The experimental work presented in Part I also explored the impact of defusion on material that varied in personal relevance. Specifically, Experiment 1 employed generic self-statements, while Experiment 2 employed self-generated statements. It remains difficult to determine the effect of this variable on the data. For example, some of the negative self-statements employed in Experiment 1 could have been highly personally relevant for some individuals, whereas other statements may have been less so. Nevertheless, the results overall suggest that defusion operates equally well with all types of negative psychological content in terms of decreases in discomfort and believability, although further research is necessary.

#### *Defusion in Relation to Thought Control and Placebo*

Similar to the experiments reported by Masuda et al. (2004), Experiment 2 of the current thesis investigated the impact of defusion relative to other strategies, primarily thought control. Overall, both of these studies reported broadly similar results. Specifically, both reported reductions in the emotional impact of negative self-referents following defusion and thought control, but defusion produced the largest reductions overall. In spite of the similar outcomes, participants in the Masuda Thought Control Condition were presented with a number of possible methods to control their thoughts (e.g. positive self-talk, abdominal breathing, positive imagery). In contrast, all participants in Experiment 2 were presented with a single word replacement exercise. This latter manipulation was included to provide better experimental control, in the event of the possibility that no one thought control strategy had been employed by participants in the original research. While the latter

manipulation in Experiment 2 was probably better, the concordance of data suggests that participants in the Masuda study were all engaging in thought control but may simply have been doing so in different ways.

Despite the similarities in outcomes for thought control, differences emerged across studies in terms of the Placebo Conditions. In Experiment 2, Placebo involved neutral reading content unrelated to the research. In contrast, Masuda et al. employed Placebo as part of their Distraction Condition. However, participants in the latter study were not explicitly instructed to try to distract themselves from any features of the experiment, rather they were simply instructed to read neutral content. Furthermore, Masuda's Distraction Condition may or may not have been preceded by Defusion (due to the nature of the alternating treatment design). And in Experiment 2, Placebo may or may not have been combined with Defusion. The former study reported that distraction decreased discomfort and believability, while the latter reported a decrease in believability for Placebo instructions, only when combined with a Defusion exercise. More importantly, the double placebo condition (Placebo/Placebo) did not produce any significant decreases in levels of discomfort or believability, thus suggesting that the previous effect was a result of the Defusion exercise and not the Placebo instructions. The use of an independent groups design in Experiment 2 afforded direct comparison between conditions, whereas the alternating treatment designs employed by Masuda et al. highlighted differences from one treatment design to another. Thus, the choice of design employed in Experiment 2 allowed for a more accurate account of the impact of Placebo. And the data were relatively clear, Placebo was not associated with changes in the emotional impact of the negative self-referents.

### *Adherence Measures*

Part I of the current thesis highlights the importance of including adherence measures in clinical analogue studies. Adherence to experimental instructions and strategy is a complex issue and is often hard to balance against other simple experimental concerns such as, how easily participants can interpret what they have to do and how long they are willing to spend doing the experiment. Addressing adherence is no simple matter because it immediately raises questions about what adherence should be detected. For example, if you check early on that participants understood instructions, you probably have to check again later to see if they are following the designated strategy. It is also almost impossible to determine precisely what they did and if they did this in the way that they were advised to do it.

Experiment 1 and 2 in particular, attempted to examine adherence as a core feature of the experimental design. Participants in both Experiments 1 and 2 reported high levels of honesty in their responding. Experiment 1 also included a measure of control by defusion-related instruction, to determine the impact of the instruction on participants' emotional ratings of the self-statements. The low adherence ratings for the defusion-related instructions were entirely consistent with their lack of impact on the ratings. If participants in all three groups had not attended well to the information regarding defusion, it is unlikely that they would have been influenced by this subsequently. However, it was not the case that participants did not *understand* these instructions, because on the instruction booklet, they had been asked to tick a box to indicate that they had fully understood and all participants had done so. It also remains possible that participants interpreted the adherence question (presented at the end of the experiment) as a reference to their adherence to the main instructions regarding the presentations of the statements and the ratings. However, this seems unlikely because

the strong changes in ratings suggested no lack of adherence during this part of the experiment, whereas the limited impact of the defusion instructions does imply lack of adherence at that earlier point. Although it is difficult to provide instructions that encourage participants to attend fully to them, while at the same time avoiding them doing any experiential exercises, the current study, at the very least, highlights the importance of including adherence measures.

Experiment 2 included two further adherence measures -- difficulty of instructions and daily use strategy. Differences emerged across conditions in terms of their ratings of difficulty in following the instructions. Participants in Thought Control/Thought Control rated their instructions as the most difficult (mean of 50, where 100 equals extremely difficult). In contrast, participants presented with either a Thought Control rationale or Thought Control exercise combined with some other element (e.g. defusion or placebo) did not report exceptionally higher ratings in this regard. The data suggest therefore, that participants did not benefit from the intended internal consistency between instructions and exercises in terms of thought control. In contrast, internal consistency did work for defusion. It is possible that these results emerged in part, from individuals' previous history with thought control. For example, although the instructions may have been consistent with what participants already do, the exercise may not have been. Thus, the intended internal consistency may not have aided participants in this context. Nevertheless, it remains difficult to determine what effect, if any, the higher difficulty ratings had overall on the experimental results.

Differences also emerged in Experiment 2 on reported levels of daily use of the strategy employed. Generally, participants presented with the Thought Control strategy were significantly more likely to have used this strategy on a daily basis than those presented with the Defusion strategy. Although this may suggest that the

positive effects for the Defusion strategy emerged due to the novelty of the exercise, it should be emphasised that overall, across all nine conditions, the use of defusion as a strategy was relatively low. Even if one argued that the positive effects of defusion were due to the novelty of the exercise, this in and of itself is not necessarily problematic as defusion strategies are designed to help clients who have been engaging in unworkable strategies. Thus, encouraging participants to do something different appears to have a positive impact on one's reactions to their negative self content. In any case, it is entirely consistent with ACT that the general use of defusion is considerably lower than the use of thought control.

#### *Summary of Findings from Part II: Chapters 4, 5 and 6*

Part II of the current thesis attempted to explain the conflicting evidence obtained across studies in terms of the self-reported ratings of the impact of defusion on discomfort, believability and willingness associated with negative thoughts. Specifically, the experimental work in the second half of the thesis was designed to assess the emotional impact of defusion on *implicit* measures of self-regard in an attempt to circumvent the problems associated with self-report measures.

Chapter 4 presented three exploratory experiments, the primary aim of which was to begin to create an IRAP that would tap into implicit self-relevant content (e.g. low self-esteem) in a manner that would subsequently allow us to manipulate these implicit attitudes through defusion. All three experiments were identical in terms of the positive and negative words presented to participants as target stimuli. However, each experiment differed in terms of the type of sample stimuli presented. Specifically, Experiment 3 presented I AM versus I AM NOT, Experiment 4 presented I AM versus OTHERS ARE, while Experiment 5 presented I AM versus I SHOULD BE.

Across all three experiments, the results indicated that participants had a strong implicit positive self-regard, with little or no reaction to the comparison sample stimulus in each case. Nevertheless, the results from Experiment 5 (I AM versus I SHOULD BE) indicated the overall weakest levels of implicit positive self-regard, relative to the other two experiments.

Experiment 6 (Chapter 5) considered whether positive and negative *statements* (e.g. ‘I am so alone that it hurts’) would better reflect the type of thoughts that people experience as opposed to simple positive and negative *words* (e.g. ‘I am lonely’). The former, therefore, may be more susceptible to a defusion intervention in the context of ‘I AM versus I SHOULD BE’ IRAP. The results confirmed this hypothesis, with participants relatively less certain that they were positive when presented with whole statements as target stimuli relative to words.

Experiment 7 (Chapter 5) was designed as a preliminary exploration to determine if the Statement-IRAP would be susceptible to defusion-related changes with groups of individuals with different levels of experience with defusion. Specifically, responses were compared between a group of undergraduates (considered to have no experience with defusion) and a group of ACT therapists (considered to have considerable experience with defusion, either personally or through therapeutic work with clients). Once again, the results indicated that both groups reported a strong implicit positive self-regard. However, the ACT Group produced a relatively stronger implicit positive self-regard than the Non-ACT Group.

Experiment 8 (Chapter 6) was designed to determine the impact of a defusion intervention on implicit self-regard in a sample of individuals with no prior history of defusion. Specifically, the experiment compared Pre- and Post-Intervention IRAP responses from a group of individuals assigned to a Defusion Condition versus a

Placebo Condition. Consistent with previous experiments, the Pre-Intervention scores indicated that both groups reported a strong implicit positive self-regard, with strong reactions to the I AM statements, but relatively weak reactions to the I SHOULD BE statements. However, Post-intervention ratings suggested that those assigned to the Defusion Group reported a larger increase in implicit positive self-regard than the Placebo Group with respect to the I AM statements. Furthermore, participants assigned to Placebo also became more certain that they should be positive than those assigned to Defusion. Finally, both groups reported decreases in the emotional impact of their self-generated negative thought in terms of explicit ratings of levels of discomfort, believability and willingness.

#### *Theoretical Issues from Part II: Chapters 4, 5 and 6*

##### *The Difficulty in Generating Appropriate Stimuli*

One of the primary aims of Part II of the current research was to determine if the IRAP would provide a supplemented measure of defusion to the self-report measures previously employed. Thus, the research presented here was designed to begin to identify the types of implicit cognitions that would be most relevant for use with a defusion manipulation. Specifically, we attempted to generate an implicit task in which self-esteem was the low, so that defusion may be effective. To this extent, Experiments 3, 4 and 5 manipulated the type of sample stimuli that acted as a comparison to the I AM stimuli presented during the IRAP. The results indicated that the I AM versus I SHOULD BE manipulation (Experiment 5) produced the significantly smallest positive implicit self-regard, while the I AM versus OTHERS ARE manipulation (Experiment 4) produced the significantly largest positive implicit self-regard. The later effect is perhaps unsurprising given that previous research with

the self-esteem IAT indicated that when the ‘other’ remains unspecified, individuals may automatically evaluate the other as negative. This allows for an inflated positive self-regard relative to when the ‘other’ is specified as being positive and suggests the role of a context effect (Karpinski, 2004). Thus in the current research, asking participants to respond to how they viewed themselves in the context of ‘others’ may not have created as much uncertainty for implicit self-regard as responding in the context of the ‘ideal self’ (i.e. I SHOULD BE).

Despite these differences, it remains the case that across all IRAP experiments presented here participants reported a strong implicit positive self-regard. Perhaps this is not surprising given the nature of the population employed in the current studies (university students and ACT therapists) and the high ratings obtained on the explicit measure of self-esteem. Indeed, the effects obtained here are entirely consistent with those reported elsewhere in the literature, which has established that the majority of people usually demonstrate high self-esteem when required to provide self-evaluations -- a finding that has been observed when using both explicit and implicit measures (Greenwald et al., 2002; Greenwald, & Farnham, 2000; Koole, Dijksterhuis, & von Kippenberg, 2001). Research has even indicated high levels of positive implicit self-esteem in formally depressed patients (both before and after negative mood induction), as well as, currently depressed patients (de Raedt, Schacht, Franck, & de Houwer, 2006; Gemar, Segal, Sagrati, & Kennedy, 2001).

Despite the strong reaction produced in response to the I AM trial-types, participants produced little reaction to the comparison stimuli across the three experiments. Although we attempted to circumvent this problem by replacing the positive and negative words with whole statements (Experiment 6), it remained difficult to elicit a strong reaction to the comparison stimuli. It is not clear at this point



why participants produced such weak responses in this regard. Some researchers have argued that an inflated positive implicit self-regard may simply reflect the context of the comparison stimuli (e.g. Karpinski, 2004), and as such, the stimuli employed here may not have been the most effective comparison stimuli for use in an IRAP designed to assess self-regard. It is equally plausible that the results may simply reflect the nature of the population employed. For example, an entirely different response set may be obtained in the context of *personally relevant* target stimuli (e.g. Houben, & Wiers, 2007). In any case, the research presented in Part II of the current thesis highlights a number of difficulties associated with generating appropriate stimuli for use with defusion interventions and implicit measures.

### *The Impact of Defusion on Implicit Cognition*

In spite of the difficulties noted above, the results of Experiments 7 and 8 suggest that defusion impacts on implicit self-regard. Consider the different patterns that emerged between the ACT and Non-ACT Groups (Experiment 7), as well as differences between the Defusion and Placebo Groups (Experiment 8). Perhaps surprisingly, defusion appeared to *increase* participants' overall positive implicit self-regard. That is, participants with a history of defusion reported stronger implicit reactions to the I AM-Positive trial-types, than those with no such history (Experiment 7). Furthermore, participants exposed to a defusion intervention reported a larger increase in Post-intervention ratings to these trial-types relative to a Placebo Condition (Experiment 8). This effect is perhaps surprising given that within therapeutic contexts, the primary target for defusion exercises is problematic psychological content, which is what we had intended with the I SHOULD BE statements. The increase in positive self-regard was perhaps unexpected, given that it

is commonly thought that for defusion to work, one must become ‘fully defused’ from both positive, as well as, negative psychological content. Based on this assumption, one might expect positive self-regard to *decrease*, to indicate a decrease in believability of these thoughts. In hindsight, the increased positive self-regard obtained with the I AM trial-types is perhaps not that surprising. If defusion is viewed as promoting psychological well-being, then there is no reason to argue that one should defuse from psychological content that does not impact negatively on the individual. Viewed in this way, the results obtained here lend some support to Wilson and Murrell’s (2004) definition that defusion is relevant in the context in which psychological content interferes with an individual behaving in the direction of their personal values. In other words, it appears that it is necessary to defuse only from problematic psychological content and, as such, is consistent with the effects of defusion obtained for the explicit ratings of positive and negative self-statements in Experiment 1.

Differences in the strength of effect for Pre- to Post-Intervention ratings for the I SHOULD BE trial-types were less distinct. Little difference emerged between the ACT and Non-ACT Groups in Experiment 7. The results obtained for Experiment 8 indicated that participants in both groups became more certain that they should be positive, with a larger increase in the strength of this effect for the Placebo Group. Again, this effect was unexpected, if defusion is seen as breaking the natural verbal processes and promoting psychological health. Perhaps these results can be explained by the small effect sizes obtained for these trial-types at baseline ratings, suggesting that participants had little or no opinion on these statements. Thus, future research should determine the impact of defusion in the context of stimuli that elicit stronger

effects to the I SHOULD BE statements and should further explore the use of alternative sample stimuli.

### *The Relationship between Implicit and Explicit Measures of Defusion*

The results outlined above suggest that defusion strategies promote psychological health in terms of increasing one's positive implicit self-regard. These contrast with those reported in Experiment 1 which indicated that although defusion decreased the emotional impact for explicit ratings of the negative self-statements, it had little impact on the positive self-statements. It is difficult to determine the root of this discrepancy in the data. However, the discrepancy between the impact of an intervention on implicit versus explicit measures is consistent with existing research. For example, Grumm et al. (2008) reported significant Post-intervention increases in explicit self-esteem in clients with chronic pain. However, no such changes were recorded for implicit measures.

It remains difficult to determine whether the results presented here reflect a dual processing system where implicit and explicit attitudes are separate systems (e.g. Fazio, & Olsen, 2003; Wilson et al., 2000). It is also possible that the discrepancy simply reflect procedural differences between the two types of task (Karpinski, 2004). Indeed, research suggests that increasing the similarity between implicit and explicit test formats increases the correlations between the two measures (Payne, Burkley, & Stokes, 2008).

### *The Relationship between Implicit Cognition and Psychological Measures*

The current research obtained mixed results in terms of the relationship between the implicit and explicit measures. Overall, there were only weak to

moderate correlations between the two. Specifically, no significant correlations emerged between the implicit and explicit measures in Experiment 3, while a positive correlation was obtained between the IRAP and the RSES in Experiment 4.

A different pattern of results emerged for the experiments that employed I AM versus I SHOULD BE as the sample stimuli, depending on the nature of the target stimuli employed (words versus statements) and the level of participants' experience with defusion. Specifically, in the context of positive and negative words (Experiment 5), large negative correlations emerged between the IRAP and both the stress and depression sub-scales. However, all correlations that emerged between measures in the context of positive and negative statements were weak (Experiment 6), and this pattern was not affected by one's history with defusion techniques (Experiment 7). Despite differences in explicit levels of awareness and acceptance between the ACT and Non-ACT Groups, no differences were reported between the groups in the correlational analysis.

Yet another pattern of results emerged in Experiment 8 and this further depended on the time frame involved (i.e. Pre- or Post-Intervention). Overall, Pre-Intervention scores indicated only weak to moderate correlations between the two measures, with the exception of a large negative correlation between the implicit measure and acceptance for the Placebo Group (no correlations emerged for the Defusion Group at Pre-Intervention). Post-intervention ratings produced a different pattern again, with large negative correlations between the implicit measure and both anxiety and depression sub-scales, while a positive correlation emerged with explicit self-esteem. Interestingly, no correlations emerged for the Placebo Group at Post-Intervention).

Taken together, these results indicated no clear pattern to the relationship between implicit and explicit measures of self-regard. This is entirely consistent with findings elsewhere in the literature. For example, a large-scale study investigating implicit and explicit self-esteem reported poor correlations between the two measures (e.g. Bosson et al., 2000). As is the case with the discrepancies between implicit and explicit measures of defusion, further research is necessary to determine if the differences are driven by dual-processing systems or are attributable to procedural issues.

### *Explicit Measures*

Part II of the current thesis incorporated a number of explicit measures across all five studies. First, participants completed a number of pre-experimental psychological measures. The results from Experiment 7 indicated between group differences in this regard. As expected, the ACT Group reported significantly higher levels of acceptance and awareness relative to the Non-ACT Group. Unexpectedly, however, differences were also reported between the Defusion and Placebo Groups in Experiment 8, with those assigned to the Placebo Group reporting significantly higher levels of acceptance. It is difficult to determine the impact of these between group differences on the implicit ratings of self-regard, given the overall lack of relationship between the two measures. Nevertheless, for issues of experimental control, future studies should attempt to match groups of participants across the full range of measures.

Participants also completed a number of IRAP-related Response Scales designed to determine their explicit reactions to the positive and negative stimuli employed in the IRAP. Experiments 3, 4 and 5 simply asked participants to rate the

relevant words for levels of positivity/negativity, discomfort, believability and willingness. Nevertheless, it remained the case the participants were simply responding to the valence of each word, as opposed to relating the word to the self. Thus, Experiments 7 and 8 modified the format in which the stimuli were presented. Specifically, each item was prefixed with the phrase “I am” to indicate the specific context to which participants were required to respond to. As expected, across all experiments participants rated the positive stimuli as positive and negative stimuli as negative, thus their opinions are consistent with the experimental views of the stimuli. Furthermore, participants rated the positive stimuli as more comfortable, more believable and were more willing to experience these words relative to the negative ones. No between group differences emerged in responses on these scales. Furthermore, no correlations emerged between any of these measures and the implicit measure (with the exception of one correlation in Experiment 8 which produced a large negative correlation between explicit believability of positive statements and implicit ratings of the I SHOULD BE trial-types for the Defusion Group only).

Third, participants in Experiment 8 were required to provide explicit (discomfort, believability and willingness) ratings of their personalised self-generated thought. This format was similar to Experiment 2, except that the intervention in 8 was presented between Pre- and Post-IRAP. Interestingly, the explicit outcomes were not consistent across the two studies. In Experiment 8, Defusion and Placebo significantly decreased discomfort and believability and increased willingness. But, in Experiment 2 willingness remained unchanged in all cases. Furthermore, Defusion decreased discomfort and believability, but Placebo did not. Interestingly, the former relationship between Placebo and discomfort/believability was consistent with the findings reported originally by Masuda et al. (2004). These variations in outcomes,

therefore, lend some support to the view that explicit measures may or may not be reliable for studying this type of psychological content.

Fourth, participants in Experiment 8 also provided explicit ratings on two adherence measures: (1) honesty in providing ratings and (2) difficulty in following instructions. Consistent with Experiment 2, all participants reported high levels of honesty and low levels of difficulty associated with following the intervention instructions. Importantly, no between group differences were reported in this regard.

### *The Utility of the IRAP in the Context of Defusion*

One of the primary aims of the Part II of the current thesis was to determine the utility of the IRAP in the context of defusion. The results presented in Experiments 7 and 8 suggest that defusion does impact on implicit self-regard. Nevertheless, the differences obtained between groups across these experiments did not reach statistical significance. Although significant post-treatment changes were reported elsewhere for the IAT (Grumm et al., 2008; Teachman, & Woody, 2003), these changes followed more complete treatment regimes with clinical populations. Thus, it is difficult to determine the size of the effect that one would expect from defusion here, particularly given that no previous IRAP research exists in this area.

The lack of a significant effect here may, in part, be explained by the small effect size obtained for the I SHOULD BE trial-types. It may therefore be of significant merit to investigate the impact of defusion in sub-clinical and clinical populations who may react differently to the stimuli presented here. At the very least, the current work offers an important initial step into developing a defusion-relevant IRAP and suggests that future development in this area is of value. Furthermore, it is the first IRAP study to investigate Pre- to Post-Intervention changes and as such,

provides an important contribution to the literature. Specifically, the inclusion of an implicit IRAP measure begins to present a fuller picture of the internal and external events that coincide with mental health. It is not about whether one measure is better than the other. It is more about getting a sense of both and determining their relationship and their impact on emotion and behaviour. Ultimately, we will likely need both to determine whether therapeutic interventions are helpful.

### *Concluding Comments*

The main aims of the current thesis were: (1) to develop an understanding of the underlying processes associated with defusion and (2) to identify the contexts in which defusion works best. To this extent, the results from Part I of the thesis suggest that defusion works best when experienced in exercise format as opposed to simply receiving defusion instructions. Furthermore, defusion appears to work in a similar means irrespective of whether participants are engaging with personally selected self-relevant thoughts or self-statements generated for experimental purposes. Specifically, in all contexts participants reported a decrease in the emotional impact of negative statements – that is, decreased discomfort and believability. However, mixed results were reported with participants' willingness to experience the statements.

In addition, the results from Part II of the thesis suggest that defusion operates in different ways, depending on the type of measure used. Specifically, explicit ratings suggested that defusion *decreases* the emotional impact of negative self-statements, but had relatively little impact on positive statements. In contrast, implicit ratings suggested that defusion *increases* participant's positive self-regard but had relatively little effect on views of how they should be. With respect to the latter effect,



however, there were a number of procedural difficulties that should be addressed before firm conclusions can be made.

The work presented here provides a valuable contribution to the limited existing literature in the area of cognitive defusion. Specifically, it is the first to offer an account of the impact of defusion on both explicit and implicit measures. The data also provide the first account of the sensitivity of the IRAP to therapeutic interventions. Furthermore, the research demonstrates the importance of experimental control in the absence of the experimenter. Taken together, the results provide one of the first comprehensive experimental analyses of the underlying psychological processes associated with cognitive defusion. Overall, the research highlights the positive emotional benefits of engaging in defusion strategies.

## References

- Abramowitz, J.S., Tolin, D.F., & Street, G.P. (2001). Paradoxical effects of thought suppression: A meta-analysis of controlled studies. *Clinical Psychology Review, 21*, 683-703.
- Abramowitz, J.S., Whiteside, S., Lynam, D., & Kalsy, S. (2003). Is thought-action fusion specific to obsessive-compulsive disorder? A mediating role of negative affect. *Behaviour Research and Therapy, 41*, 1069-1079.
- American Psychiatric Association (1994). *Diagnostic and statistical manual of mental disorders* (4<sup>th</sup> edition). Washington, D.C.: Author.
- Ames, S.L., Grenard, J.L., Thush, C., Sussman, S., Wiers, R.W., & Stacy, A.W. (2007). Comparison of indirect assessments as predictors of marijuana use among at-risk adolescents. *Experimental and Clinical Psychopharmacology, 15*, 204-218.
- Aslaksen, P.M., Myrbakk, I.N., Høifødt, R.S., & Flaten, M.A. (2007). The effect of experimenter gender on autonomic responses to pain stimuli. *Pain, 129*, 260-8.
- Bach, P., & Hayes, S.C. (2002). The use of Acceptance and Commitment Therapy to prevent the rehospitalization of psychotic patients: A randomised controlled trial. *Journal of Consulting and Clinical Psychology, 70*, 1129-1139.
- Banse, R., Seise, J., & Zerbes, N. (2001). Implicit attitudes toward homosexuality: Reliability, validity, and controllability of the IAT. *Zeitschrift für Experimentelle Psychologie, 48*, 145-160.
- 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*, 169-177.

- Baron, A.S., & Banaji, M.R. (2006). The development of implicit attitudes: Evidence of race evaluations from ages 6, 10, and adulthood. *Psychological Science, 17*, 53-58.
- Beck, A.T. (1970). Cognitive therapy: Nature and relation to behavior therapy. *Behavior Therapy, 1*, 184-200.
- Beck, A.T., Rush, A.J., Shaw, B.F., & Emery, G. (1979). *Cognitive therapy of depression*. Guilford Press: New York.
- Beck, A. T., Steer, R., & Brown, G. (1996). *Beck Depression Inventory: Second Edition*. New York: Psychological Corporation.
- Beevers, C.G., & Meyer, B. (2004). Thought suppression and depression risk. *Cognition and Emotion, 18*, 859-867.
- Benson, H., Alexander, S., & Feldman, C.L. (1975). Decreased premature ventricular contractions through use of the relaxation response in patients with stable ischemic heart disease. *Lancet, 2*, 380-382.
- Benson, H., & Klipper, M.Z. (1976). *The relaxation response*. New York: Avon Books.
- 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.
- Blascovich, J., & Tomaka, J. (1993). Measures of Self-Esteem. In: J.P. Robinson, P.R. Shaver, & L.S.Wrightsman (Eds.), *Measures of personality and social psychological attitudes*. (3<sup>rd</sup> Ed.). Ann Arbor: Institute for Social Research. (pp 115-160).

- Bond, F.W., & Bunce, D. (2000). Mediators of change in emotion-focused and problem-focused worksite stress management interventions. *Journal of Occupational Health Psychology, 5*, 156-163.
- Bond, F.W., & Bunce, D. (2003). The role of acceptance and job control in mental health, job satisfaction and work performance. *Journal of Applied Psychology, 88*, 1057-1067.
- Bond, F.W., Hayes, S.C., Baer, R.A., Carpenter, K.M., Orcutt, H.K., Waltz, T., et al. (2008). *Preliminary psychometric properties of the Acceptance and Action Questionnaire-II: A revised measure of psychological flexibility and acceptance*. Manuscript submitted for publication.
- Borkovec, T.D., & Hu, S. (1990). The effect of worry on cardiovascular response to phobic imagery. *Behaviour Research and Therapy, 28*, 69-73.
- Bosson, J.K., Swann, W.B., & Pennebaker, J.W. (2000). Stalking the perfect measure of implicit self-esteem: The blind men and the elephant revisited? *Journal of Personality and Social Psychology, 79*, 631-643.
- Campbell-Sills, L., Barlow, D.H., Brown, T.A., & Hofmann, S.G. (2006). Effects of suppression and acceptance on emotional responses of individuals with anxiety and mood disorders. *Behaviour Research and Therapy, 44*, 1251-1263.
- Cardaciotto, L., Herbert, J.D., Forman, E.M., Moitra, E., & Farrow, V. (in press). The assessment of present-moment awareness and acceptance: The Philadelphia Mindfulness Scales. *Assessment*.
- Cioffi, D., & Holloway, J. (1993). Delayed costs of suppressed pain. *Journal of Personality and Social Psychology, 64*, 274-282.

- Clark, D.M., Ball, S., & Pape, D. (1991). An experimental investigation of thought suppression. *Behaviour Research and Therapy*, *29*, 253-257.
- Cohene, L.S., Smith, M.C., & Klein, D. (1978). Semantic satiation revisited with a lexical decision task. *Memory and Cognition*, *6*, 131-140.
- Cordova, J.V., Jacobsen, N.S., & Christensen, A. (1998). Acceptance versus change in behavioral couples therapy: Impact on client communication processes in the therapy session. *Journal of Marital and Family Therapy*, *24*, 437-455.
- Craske, M.G., Miller, P.P., Rotunda, R., & Barlow, D.H. (1990). A descriptive report of features of initial unexpected panic attacks in minimal and extensive avoiders. *Behaviour Research and Therapy*, *28*, 395-400.
- Cullen, C., & Barnes-Homes, D. (in press). Implicit pride and prejudice: A heterosexual phenomenon? In T.G. Morrison & M.A. Morrison (Eds), *Modern prejudice*. New York: Nova Science.
- Cunningham, W.A., Johnson, M.K., Raye, C.L., Gatenby, J.C., Gore, J.C., & Banaji, M.R. (2004). Separable neural components in the processing of Black and White faces. *Psychological Science*, *15*, 806-813.
- Dahl, J., Wilson, K., & Nillson, A. (2004). Acceptance and Commitment Therapy and the treatment of persons at risk for long-term disability resulting from stress and pain symptoms: A preliminary randomised trial. *Behavior Therapy*, *35*, 785-801.
- Dasgupta, N., McGhee, D.E., Greenwald, A.G., & Banaji, M.R. (2000). Automatic preference for White Americans: Eliminating the familiarity explanation. *Journal of Experimental Social Psychology*, *36*, 316-328.

- 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 Houwer, J. (2003). The Extrinsic Affective Simon Task. *Experimental Psychology, 50*, 77-85.
- de Jong, P.J. (2002). Implicit self-esteem and social anxiety: Differential self-favouring effects in high and low anxious individuals. *Behaviour Research and Therapy, 40*, 501-508.
- de Jong, P.J. Passman, W., Kindt, M., & van den Hout, M.A. (2001). A reaction time paradigm to assess (implicit) complaint-specific dysfunctional beliefs. *Behaviour Research and Therapy, 39*, 101-113.
- de Raedt, R., Schacht, R., Franck, E., & de Houwer, J. (2006). Self-esteem and depression revisited: Implicit positive self-esteem in depressed patients? *Behaviour Research and Therapy, 44*, 1017-1028.
- Donaldson-Feilder, E.J., & Bond, F.W. (2004). The relative importance of psychological acceptance and emotional intelligence to workplace well-being. *British Journal of Guidance and Counselling, 32*, 187-203.
- Egloff, B., & Schmuckle, S.C. (2002). Predictive validity of an Implicit Association Test for assessing anxiety. *Journal of Personality and Social Psychology, 83*, 1441-1455.
- Ehlers, A., & Steil, R. (1995). Maintenance of intrusive memories in posttraumatic stress disorder: A cognitive approach. *Behavioural and Cognitive Psychotherapy, 23*, 217-249.
- Esposito, N.J., & Phelton, L.H. (1971). Review of the measurement of semantic Satiation. *Psychological Bulletin, 75*, 330-346.

- Fazio, R.H., Jackson, J.R., Dunton, B.C., & Williams, C.J. (1995). Variability in automatic activation as an unobtrusive measure of racial attitudes: A bona fide pipeline? *Journal of Personality and Social Psychology*, *69*, 1013-1027.
- Fazio, R.H., & Olson, M.A. (2003). Implicit measures in social cognition research: Their meaning and use. *Annual Review of Psychology*, *54*, 297-327.
- Fazio, R.H., Sanbonmatsu, D.M., Powell, M.C., & Kardes, F.R. (1986). On the automatic activation of attitudes. *Journal of Personality and Social Psychology*, *50*, 229-238.
- Feldner, M.T., Zvolensky, M.J., Eifert, G.H., & Spira, A.P. (2003). Emotional avoidance: An experimental test of individual differences and response suppression using biological challenge. *Behaviour Research and Therapy*, *41*, 403-411.
- Fernandez, E., & Turk, D.C. (1994). Demand characteristics underlying ratings of sensory versus affective components of pain. *Journal of Behavioural Medicine*, *17*, 375-390.
- Fiedler, K., & Bluemke, M. (2005). Faking the IAT: Aided and unaided response control on the Implicit Association Test. *Basic and Applied Social Psychology*, *27*, 307-316.
- Field, A. P., & Lawson, J. (2003). Fear information and the development of fears during childhood: Effects on implicit fear responses and behavioural avoidance. *Behaviour Research and Therapy*, *41*, 1277-1293.
- Gaillard, V., Vandenberghe, M., Destrebecqz, A., & Cleeremans, A. (2006). First- and third-person approaches in implicit learning research. *Consciousness and Cognition*, *15*, 709-722.

- Gamer, J., Schmukle, S.C., Luka-Krausgrill, U., & Egloff, B. (in press). Examining the dynamics of the implicit and explicit self-concept in social anxiety: Changes in the implicit assessment test-anxiety and the social phobia anxiety inventory. *Journal of Personality Assessment*.
- Gaudiano, B.A., & Herbert, J.D. (2006). Acute treatment of inpatients with psychotic symptoms using Acceptance and Commitment Therapy: Pilot results. *Behaviour Research and Therapy*, *44*, 415-437.
- Gawronski, B., Hofman, W., & Wilbur, C.J. (2006). Are “implicit” attitudes unconscious? *Consciousness and Cognition*, *15*, 485-499.
- Gemar, M.C., Segal, Z.V., Sagrati, S., & Kennedy, S.J. (2001). Mood-induced changes on the Implicit Association Test in recovered depressed patients, *Journal of Abnormal Psychology*, *110*, 282-289.
- Gifford, E.V., Kohlenberg, B.S., Hayes, S.C., Antonuccio, D.O., Piasecki, M.M., Rasmussen-Hall, M.L. et al. (2004). Acceptance-based treatment for smoking cessation. *Behavior Therapy*, *35*, 689-705.
- Gilboa-Schechtman, E., & Foa, E.B. (2001). Patterns of recovery from trauma: The use of intraindividual analysis. *Journal of Abnormal Psychology*, *110*, 392-400.
- Gratz, K.L., Bornovalova, M.A., Delany-Brumsey, A., Nick, B., & Lejuez, C.W. (2007). A laboratory-based study of the relationship between childhood abuse and experiential avoidance among inner-city substance users: The role of emotional nonacceptance. *Behavior Therapy*, *38*, 256-268.
- Greenberger, E., Chen, C., Dmitrieva, J., & Farruggia, S.P. (2003). Item-wording and the dimensionality of the Rosenberg Self-Esteem Scale: Do they matter? *Personality and Individual Differences*, *35*, 1241-1254.



- Greenwald, A.G. (1990). What cognitive representations underlie social attitudes? *Bulletin of the Psychonomic Society*, 28, 254-260.
- Greenwald, A.G., & Banaji, M.R. (1995). Implicit social cognition: Attitudes, self-esteem, and stereotypes. *Psychological Review*, 102, 4-27.
- Greenwald, A.G., Banaji, M.R., Rudman, L.A., Farnham, S.D., Nosek, B.A., & Mellot, D.S. (2002). A unified theory of implicit attitudes, stereotypes, self-esteem and self-concept. *Psychological Review*, 109, 3-25.
- Greenwald, A.G., & Farnham, S.D. (2000). Using the Implicit Association Test to measure self-esteem and self-concept. *Journal of Personality and Social Psychology*, 79, 1022-1038.
- 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, 1464-1480.
- Greenwald, A.G., & Nosek, B.A. (2001). Health of the Implicit Association Test at age 3. *Zeitschrift für Experimentelle Psychologie*, 48, 85-93.
- Greenwald, A.G., Nosek, B.A., & Banaji, M.R. (2003). Understanding and using the Implicit Association Test: I. An improved scoring algorithm. *Journal of Personality and Social Psychology*, 85, 197-216.
- Greenwald, A.G., Nosek, B.A., & Sriram, N. (2006). Consequential validity of the Implicit Association Test: Comment on the article by Blanton and Jaccard. *American Psychologist*, 61, 56-61.
- Gregg, J.A., Callaghan, G.M., Hayes, S.C., & Glenn-Lawson, J.L. (2007). Improving diabetes self-management through acceptance, mindfulness, and values: A randomized control trial. *Journal of Consulting and Clinical Psychology*, 75, 336-343.

- Gronwall, D.M. (1977). Paced auditory serial-addition task: A measure of recovery from concussion. *Perceptual and Motor Skills*, 44, 367-373.
- Grumm, M., Erbe, K., von Collani, G., & Nestler, S. (2008). Automatic processing of pain: The change of implicit pain associations after psychotherapy. *Behaviour Research and Therapy*, 46, 701-714.
- Gutierrez, O., Luciano, C., Rodriguez, M., & Fink, B.C. (2004). Comparison between an acceptance-based and a cognitive-control-based protocol for coping with pain. *Behavior Therapy*, 35, 767-784.
- Hastings, R.P. (1997). Measuring staff perceptions of challenging behaviour: The challenging behaviour attributions scale (CHABA). *Journal of Intellectual Disability Research*, 41, 495-501.
- Hayes, S.C., Barnes-Holmes, D., & Roche, B. (2001). *Relational Frame Theory: A Post-Skinnerian account of human language and cognition*. New York: Kluwer Academic/Plenum Publishers.
- Hayes, S.C., Bisset, R.T., Korn, Z., Zettle, R.D., Rosenfard, I.S., Cooper, L.D. et al. (1999). The impact of acceptance versus control rationales on pain tolerance. *The Psychological Record*, 49, 33-47.
- 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.
- Hayes, S.C., Strosahl, K.D., & Wilson, K.G. (1999). *Acceptance and Commitment Therapy: An experiential approach to behavior change*. New York: Guilford Press

- Hayes, S.C., Strosahl, K., Wilson, K.G., Bisset, R.T., Pistorello, J., Toarmino, D. et al., (2004). Measuring experiential avoidance: A preliminary test of a working model. *The Psychological Record*, *54*, 553-578.
- Hayes, S.C., Wilson, K.G., Gifford, E.V., Bisset, R., Piasecki, M., Batten, S.V., et al., (2004). A preliminary trial of twelve-step facilitation and Acceptance and Commitment Therapy with polysubstance-abusing methadone-maintained opiate addicts. *Behavior Therapy*, *35*, 667-688.
- Hayes, S.C., Wilson, K.G., Gifford, E.V., Follette, V.M., & Strosahl, K.D. (1996). Experiential avoidance and behavioral disorders: A functional dimensional approach to diagnosis and treatment. *Journal of Consulting and Clinical Psychology*, *64*, 1152-1168.
- Hofman, W., Gawronski, B., Gschwendner, T., Le, H., & Schmitt, M. (2005). A meta-analysis on the correlation between the Implicit Association Test and explicit self-report measure. *Personality and Social Psychology Bulletin*, *31*, 1369-1385.
- Houben, K., & Wiers, R.W. (2007). Personalizing the alcohol-IAT with individualized stimuli: Relationship with drinking behavior and drinking-related problems. *Addictive Behaviors*, *32*, 2852-2864.
- Kabat-Zinn, J. (1994). *Wherever you go, there you are*. New York: Hyperion.
- Kanter, J.W., Kohlenberg, R.J., & Loftus, E.F. (2004). Experimental and psychotherapeutic demand characteristics and the cognitive therapy rationale: An analogue study. *Cognitive Therapy and Research*, *28*, 229-239.
- Karekla, M., Forsyth, J.P., & Kelly, M.M. (2004). Emotional avoidance and panicogenic responding to a biological challenge procedure. *Behavior Therapy*, *35*, 725-746.

- Karpinski, A. (2004). Measuring self-esteem using the Implicit Association Test: The role of the other. *Personality and Social Psychology Bulletin*, *30*, 22-34.
- Kashdan, T.B., Barrios, V., Forsyth, J.P., & Steger, M.F. (2006). Experiential avoidance as a generalized psychological vulnerability: Comparison with coping and emotion regulation strategies. *Behaviour Research and Therapy*, *44*, 1301-1320.
- Kessler, R.C., Berglund, P., Demler, O., Jin, R., & Walters, E.E. (2005). Lifetime prevalence and age-of-onset distributions of *DSM-IV* disorders in the national comorbidity survey replication. *Archives of General Psychiatry*, *62*, 593-602.
- Kohlenberg, R.J., & Tsai, M. (1991). *Functional analytic psychotherapy: Creating intense and curative therapeutic relationships*. New York: Plenum Press.
- Koole, S.L., Dijksterhuis, A., & von Knippenberg, A. (2001). Whats in a name: Implicit self-esteem and the automatic self. *Journal of Personality and Social Psychology*, *80*, 669-685.
- Kross, E., Ayduk, O., & Mischel, W. (2005). When asking “why” does not hurt: Distinguishing Rumination from reflective processing of negative emotions. *Psychological Science*, *16*, 709-715.
- Leary, M.R., & Kowalski, R.M. (1990). Impression management: A literature review and two-component model. *Psychological Bulletin*, *107*, 34-47.
- Leitenberg, H., Greenwald, E., & Cado, S. (1992). A retrospective study of long-term methods of coping with having been sexually abused during childhood. *Child Abuse and Neglect*, *16*, 399-407.

- Levitt, J.T., Brown, T.A., Orsillo, S.M., & Barlow, D.H. (2004). The effects of acceptance versus suppression of emotion on subjective and psychophysiological response to carbon dioxide challenge in patients with panic disorder. *Behavior Therapy, 35*, 747-766.
- Linehan, M.M. (1993). *Cognitive-behavioral treatment of borderline personality disorder*. New York: Guilford.
- Lovibond, P.F., & Lovibond, S.H. (1995). *Manual for the Depression, Anxiety and Stress Scales* (2<sup>nd</sup> Ed). Sydney: Psychological Foundation.
- Luciano, C., & Hayes, S.C. (2001). Experiential Avoidance Disorder. *International Journal of Clinical and Health Psychology, 1*, 109-157.
- Luciano, C., Rodríguez, M., & Gutiérrez, O. (2004). A proposal for synthesizing verbal contexts in Experiential Avoidance Disorder and Acceptance and Commitment Therapy. *International Journal of Psychology and Psychological Therapy, 4*, 377-394.
- Lundgren, T., Dahl, J., Melin, L., & Kies, B. (2006). Evaluation of Acceptance and Commitment Therapy for drug refractory epilepsy: A randomized control trial in South Africa – a pilot study. *Epilepsia, 47*, 2173-2179.
- Marcks, B.A., & Woods, D.W. (2005). A comparison of thought suppression to an acceptance-based technique in the management of personal intrusive thoughts: A controlled evaluation. *Behaviour Research and Therapy, 43*, 433-445.
- Marx, B.P., & Sloan, D.M. (2002). The role of emotion in the psychological functioning of adult survivors of childhood sexual abuse. *Behavior Therapy, 33*, 563-577.

- 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.
- McConnell, A.R., & Leibold, J.M. (2001). Relations between the Implicit Association Test, explicit racial attitudes, and discriminatory behavior. *Journal of Experimental Social Psychology*, *37*, 435-442.
- McCracken, L.M. (1998). Learning to live with pain: Acceptance of pain predicts adjustments in persons with chronic pain. *Pain*, *74*, 21-27.
- McCracken, L.M., & Eccleston, C. (2003). Coping or acceptance: What to do about chronic pain? *Pain*, *105*, 197-204.
- McCracken, L.M., Mackichan, F., & Eccleston, C. (2007). Contextual cognitive behavioral therapy for severely disabled chronic pain sufferers: Effectiveness and clinically significant change. *European Journal of Pain*, *11*, 314-322.
- Nahmias, A. (2002). Verbal reports on the contents of consciousness – reconsidering introspectionist methodology, *Psyche*, *8* (21).
- Najmi, S., Wegner, D.M., & Nock, M.K. (2007). Thought suppression and self-injurious thoughts and behaviours. *Behaviour Research and Therapy*, *45*, 1957-1965.
- Nisbett, R.E., & Wilson, T.D. (1977). Telling more than we can know: Verbal reports on mental processes. *Psychological Review*, *84*, 231-259.
- Nock, M.K., & Banaji, M.R. (2007). Prediction of suicide ideation and attempts among adolescents using a brief performance-based test. *Journal of Consulting and Clinical Psychology*, *75*, 707-715.
- Nosek, B.A. (2005). Moderators of the relationship between implicit and explicit evaluation. *Journal of Experimental Psychology: General*, *134*, 565-584.

- Nosek, B.A. (2007). Implicit-explicit relations. *Current Directions in Psychological Science, 16*, 65-69.
- Nosek, B.A., & Banaji, M.R. (2001). The go/no-go association task. *Social Cognition, 19*, 625-664.
- Nosek, B.A., & Smyth, F.L. (2007). A multitrait-multimethod validation of the Implicit Association Test: Implicit and explicit attitudes are related but distinct constructs. *Experimental Psychology, 54*, 14-29.
- Nuttin, J.M. (1987). Affective consequences of mere ownership: The name letter effect in twelve European languages. *European Journal of Social Psychology, 17*, 381-402.
- Orsillo, S.M., & Batten, S.V. (2005). Acceptance and Commitment Therapy in the treatment of posttraumatic stress disorder. *Behavior Modification, 29*, 95-129.
- Orsillo, S.M., Roemer, L., Block Lerner, J., & Tull, T.M. (2004). Acceptance, mindfulness, and cognitive-behavioral therapy: Comparisons, contrasts, and application to anxiety. In S.C. Hayes, V.M. Follette, & M.M. Linehan (Eds.), *Mindfulness and acceptance: Expanding the cognitive-behavioral tradition* (pp. 66-95). New York: Guilford.
- Paulhus, D.L. (1988). Assessing self deception and impression management in self-reports: The balanced inventory of desirable responding.
- Paulhaus, D.L. (1991). Measurement and control of response bias. In J.P. Robinson, P.R. Shaver, & L.S. Wrightsman (Eds.), *Measures of personality and social psychological attitudes* (pp. 17-59). New York: Academic Press.
- Payne, B.K., Burkley, M.A., & Stokes, M.B. (2008). Why do implicit and explicit attitude tests diverge? The role of structural fit. *Journal of Personality and Social Psychology, 94*, 16-31.

- Plumb, J.C., Orsillo, S.M., & Luterek, J.A. (2004). A preliminary test of the role of experiential avoidance in post-event functioning. *Journal of Behavior Therapy and Experimental Psychiatry, 35*, 245-257.
- Polusny, M.A., Rosenthal, M.Z., Aban, I., & Follette, V.M. (2004). Experiential avoidance as a mediator of the effects of adolescent sexual victimization on negative adult outcomes. *Violence and Victims, 19*, 109-120.
- Pronin, E., & Krugler, M.B. (2007). Valuing thoughts, ignoring behavior: The introspection illusion as a source of the bias blind spot. *Journal of Experimental Social Psychology, 43*, 565-578.
- Purdon, C. (1999). Thought suppression and psychopathology. *Behaviour Research and Therapy, 37*, 1029-1054.
- Purdon, C., & Clark, D.A. (1994). Perceived control and appraisal of obsessional intrusive thoughts: A replication and extension. *Behavioural and Cognitive Psychotherapy, 22*, 26-285.
- Purdon, C., Rowa, K., & Antony, M.M. (2005). Thought suppression and its effects on thought frequency, appraisal and mood state in individuals with obsessive-compulsive disorder. *Behaviour Research and Therapy, 43*, 93-108.
- Rassin, E., Diepstraten, P., Merckelbach, H., & Muris, P. (2001). Thought-action fusion and thought suppression in obsessive-compulsive disorder. *Behaviour Research and Therapy, 39*, 757-764.
- Rassin, E., Merckelbach, H., Muris, P., & Schmidt, H. (2001). The thought-action fusion scale: Further evidence for its reliability and validity. *Behaviour Research and Therapy, 39*, 537-544.



- Robinson, M.D., Meier, B.P., Zetocha, K.J., & McCaul, K.D. (2005). Smoking and the Implicit Association Test: When the contrast category determines the theoretical conclusions. *Basic and Applied Social Psychology, 27*, 201-212.
- Roemer, L., & Borkovec, T.D. (1994). Effects of suppressing thoughts about emotional material. *Journal of Abnormal Psychology, 103*, 467-474.
- Rosenberg, M. (1965). *Society and the adolescent self-image*. Princeton, NJ: Princeton University Press.
- Rosenberg, M.J. (1969). The conditions and consequences of evaluation apprehension. In R. Rosenthal & R.L. Rosnow (Eds.), *Artifact in behavioral research* (pp. 279-349). New York: Academic Press.
- Rosenthal, M.Z., Cheavens, J.S., Lynch, T.R., & Follette, V.M. (2006). Thought suppression mediates the relationship between negative mood and PTSD symptom severity in sexually assaulted women. *Journal of Traumatic Stress, 19*, 741-745.
- Salkovskis, P.M., & Campbell, P. (1994). Thought suppression induces intrusion in naturally occurring negative intrusive thoughts. *Behaviour Research and Therapy, 32*, 1-8.
- Salters-Pedneault, K., Tull, M.T., & Roemer, L. (2004). The role of avoidance of emotional material in the anxiety disorders. *Applied and Preventive Psychology, 11*, 95-114.
- Sanchez-Craig, M. (1984). *A therapists' manual for secondary prevention of alcohol problems: Procedures for teaching moderate drinking and abstinence*. Toronto: Addiction Research Foundation.

- Schmukle, S.C., & Egloff, B. (2004). Does the Implicit Association Test for assessing anxiety measure trait or state variance? *European Journal of Personality, 18*, 483-494.
- Segal, Z.V., Williams, J.M.G., & Teasdale, J.D. (2002). *Mindfulness-based cognitive therapy for depression: A new approach to preventing relapse*. New York: Guilford Press.
- Shafran, R., & Rachman, S. (2004). Thought-action fusion: A review. *Journal of Behavior Therapy and Experimental Psychology, 35*, 87-107.
- Shafran, R., Thordarson, D., & Rachman, S. (1996). Thought-action fusion in obsessive compulsive disorder. *Journal of Anxiety Disorders, 10*, 379-391.
- Smith, L.C. (1984). Semantic satiation affects category membership decision time but not lexical priming. *Memory and Cognition, 12*, 483-488.
- Spielberger, C.D., Gorsuch, R.L., Lushene, R., Vagg, P.R., & Jacobs, G.A. (1983). *Manual for the State-Trait Anxiety Inventory*. California: Consulting Psychologists Press.
- Stewart, S.H., Zvolensky, M.J., & Eifert, G.H. (2002). The relations of anxiety sensitivity, experiential avoidance, and alexithymic coping to young adults' motivations for drinking. *Behavior Modification, 26*, 274-296.
- Takahashi, M., Muto, T., Tada, M., & Sugiyama, M. (2002). Acceptance rationale and increasing pain-tolerance: Acceptance-based and FEAR-based practice. *Japanese Journal of Behavior Therapy, 28*, 35-46.
- Teachman, B.A., Gapinski, K.D., Brownell, K.D., Rawlins, M., & Jeyaram, S. (2003). Demonstrations of implicit anti-fat bias: The impact of providing causal information and evoking empathy. *Health Psychology, 22*, 68-78.

- Teachman, B.A., Gregg, A.P., & Woody, S.R. (2001). Implicit associations for fear-relevant stimuli among individuals with snake and spider fears. *Journal of Abnormal Psychology, 110*, 226-235.
- Teachman, B.A., & Woody, S.R. (2003). Automatic processing in spider phobia: Implicit fear associations over the course of treatments. *Journal of Abnormal Psychology, 112*, 100-109.
- Teasdale, J.D., Segal, Z.V., Williams, J.M.G., Ridgeway, V.A., Soulsby, J.M., & Lau, M.A. (2000). Prevention of relapse/recurrence in major depression by mindfulness-based cognitive therapy. *Journal of Consulting and Clinical Psychology, 68*, 615-623.
- The Psychological Society of Ireland. (2003). *Code of professional ethics*. Dublin: Author.
- Thush, C., & Wiers, R.W. (2007). Explicit and implicit alcohol-related cognitions and the prediction of future drinking in adolescents. *Addictive Behaviors, 32*, 1367-1383.
- Titchener, E.B. (1916). *A text-book of psychology*. New York: MacMillan.
- Tull, M.T., & Gratz, K.L. (2008). Further examination of the relationship between anxiety sensitivity and depression: The mediating role of experiential avoidance and difficulties in engaging in goal-directed behaviour when distressed. *Journal of Anxiety Disorders, 22*, 199-210.
- Tull, M.T., & Roemer, L. (2007). Emotion regulation difficulties associated with the experience of uncued panic attacks: Evidence of experiential avoidance, emotional nonacceptance and decreased emotional clarity. *Behavior Therapy, 38*, 378-391.

- Vahey, N.A., Barnes-Holmes, D., Barnes-Holmes, Y., & Stewart, I. (in press). A first test of the Implicit Relational Assessment Procedure (IRAP) as a measure of self-esteem: Irish prisoner groups and university students. *The Psychological Record*.
- Watkins, E.R., & Moulds, M. (2007). Revealing negative thinking in recovered major depression: A preliminary investigation. *Behaviour Research and Therapy*, *45*, 3069-3076.
- Webster, D. (2002, January). China's unknown Gobi: Alashan. *National Geographic*, *201*, 48-75.
- Wegner, D.M. (1994). Ironic processes of mental control. *Psychological Review*, *101*, 34-52.
- Wenger, D.M., & Erber, R. (1992). The hyperaccessibility of suppressed thoughts. *Journal of Personality and Social Psychology*, *63*, 903-912.
- Wegner, D.M., Erber, R., & Zanakos, S. (1993). Ironic processes in the mental control of mood and mood-related thought. *Journal of Personality and Social Psychology*, *65*, 1093-1104.
- Wegner, D.M., Schneider, D.J., Carter, S.R., & White, T.L. (1987). Paradoxical effects of thought suppression. *Journal of Personality and Social Psychology*, *53*, 5-13.
- Wegner, D.M., Schreiber, D.J., Knutson, B., & McMahon, S.R. (1991). Polluting the stream of consciousness: The effect of thought suppression of the mind's environment. *Cognitive Therapy and Research*, *15*, 141-151.
- Wegner, D.M., & Zanakos, S. (1994). Chronic thought suppression. *Journal of Personality*, *62*, 615-640.

- Wells, A. (2000). *Emotional disorders and metacognition: Innovative cognitive therapy*. Chichester, UK: Wiley
- Wenzlaff, R.M., & Wegner, D.M. (2000). Thought suppression. *Annual Review of Psychology, 51*, 59-91.
- Wenzlaff, R.M., Wegner, D.M., & Klein, S.B. (1991). The role of thought suppression in the bonding of thought and mood. *Journal of Personality and Social Psychology, 60*, 500-508.
- Wenzlaff, R.M., Wegner, D.M., & Roper, D.W. (1998). Depression and mental control: The resurgence of unwanted negative thoughts. *Journal of Personality and Social Psychology, 55*, 882-892.
- Wicksell, R.K., Renöfält, J., Olsson, G.L., Bond, F.W., & Melin, L. (2008). Avoidance and cognitive fusion – Central components in pain related disability? Development and preliminary validation of the Psychological Inflexibility in Pain Scale (PIPS). *European Journal of Pain, 12*, 491-500.
- Wiers, R.W., Houben, K., & de Kraker, J. (2007). Implicit cocaine associations in active cocaine users and controls. *Addictive Behaviors, 32*, 1284-1289.
- Williams, J.M.G., Mathews, A., & McCleod, C. (1996). The emotional stroop task and psychopathology. *Psychological Bulletin, 120*, 3-24.
- Wilson, T.D., Lindsey, S., & Schooler, T.Y. (2000). A model of dual attitudes. *Psychological Review, 107*, 101-126.
- Wilson, K.G., & Murrell, A.R. (2004). Values work in Acceptance and Commitment Therapy: Setting a course for behavioral treatments. In S.C. Hayes, V.M. Follette, & M.M. Linehan (Eds.) *Mindfulness and acceptance: Expanding the cognitive-behavioral tradition* (pp.120-151). New York: Guilford.

- Woods, D. W., Wetterneck, C.T., & Flessner, C.A. (2006). A controlled evaluation of Acceptance and Commitment Therapy plus habit reversal for trichotillomania. *Behaviour Research and Therapy*, *44*, 639-656.
- Zettle, R.D. (2003). Acceptance and Commitment Therapy versus systematic desensitization in treatment of mathematics anxiety. *The Psychological Record*, *53*, 197-215.
- Zettle, R.D. (2007). *ACT for depression: A clinician's guide to using Acceptance and Commitment Therapy in treating depression*. California: New Harbinger.
- Zettle, R.D., & Hayes, S.C. (1986). Dysfunctional control by client verbal behavior: The context of reason-giving. *The Analysis of Verbal Behavior*, *4*, 30-38.

## Appendices

- Appendix A The Acceptance and Action Questionnaire (AAQ-35)
- Appendix B The Beck Depression Inventory (BDI-II)
- Appendix C The State Trait Anxiety Inventory (STAI Form Y-2)
- Appendix D Mean Ratings for the Forty Self-statements (Twenty Positive and Twenty Negative) used in Experiment 1
- Appendix E Consent Form Employed in Experiment 1
- Appendix F The Acceptance and Action Questionnaire (AAQ-49)
- Appendix G The Depression, Anxiety and Stress Scales (DASS-21)
- Appendix H The Balanced Inventory of Desirable Responding (BIDR)
- Appendix I The Rosenberg Self-Esteem Scale (RSES)
- Appendix J The Acceptance and Action Questionnaire-II (AAQ-II)
- Appendix K An Example of an IRAP-Related Response Scale for Experiments 3, 4 and 5
- Appendix L The Philadelphia Mindfulness Scales (PHLMS)
- Appendix M Some Examples from the IRAP-Related Believability Scale for Experiments 7 and 8
- Appendix N Experimental Guide to Complete the Online IRAP in Experiment 7

## Appendix A

### The Acceptance and Action Questionnaire (AAQ-35)

Below you will find a list of statements. Please rate the truth of each statement as it applies to you. Use the following scale to make your choice.

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

1. Thoughts can be dangerous.	Reverse
2. I do not try to change my unhappy, or fearful, thoughts.	
3. I can't do things, if I am unhappy.	Reverse
4. It is more important to move towards my goals, than to feel good.	
5. My thoughts and feelings can get in the way of my success.	Reverse
6. There is nothing wrong with having unhappy thoughts and feelings.	
7. I try to achieve my goals, even if I am uncertain that I can.	
8. If I feel fearful, then there is really something to be fearful about.	Reverse
9. If I feel unhappy, then I should determine the reason for my unhappiness.	Reverse
10. If I value something, I'll work for it, even if I disappoint people by doing so.	
11. I choose to get on with my life, rather than struggle with my worries or unhappiness.	
12. The greater my worries or anxieties become, the more concerned I get for myself.	Reverse
13. Even if I fear I may get it wrong, I can still take action on a problem.	
14. I should act according to my feelings at the time.	Reverse
15. I try to suppress thoughts and feelings that I don't like by just not thinking about them.	Reverse
16. It's OK to feel depressed or anxious.	



### Appendix A (continued)

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

17. Despite doubts, I feel as though I can set a course in my life and then stick to it.	
18. If I could magically remove all the painful experiences I've had in my life, I would do so.	Reverse
19. I worry about not being able to control my anxieties, worries, and feelings.	Reverse
20. Anxiety is bad.	Reverse
21. I'm not afraid of my feelings.	
22. I am in control of my life.	
23. If I get bored of a task, I can still complete it.	
24. Worries can get in the way of my success.	Reverse
25. I am able to take action on a problem even if I am uncertain what is the right thing to do.	
26. If I promised to do something, I'll do it, even if I later don't feel like it.	
27. When I feel depressed or anxious, I am unable to take care of my responsibilities.	Reverse
28. I try hard to avoid feeling depressed or anxious.	Reverse
29. I often catch myself daydreaming about things I've done and what I would do differently next time.	Reverse
30. I rarely worry about getting my anxieties, worries, and feelings under control.	
31. When I evaluate something negatively, I usually recognize that this is just a reaction, not an objective fact.	
32. When I compare myself to other people, it seems that most of them are handling their lives better than I do.	Reverse

**Appendix A (continued)**

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

33. I am able to take action on a problem even if I am uncertain what is the right thing to do.	
34. It's unnecessary for me to learn to control my feelings, in order to handle my life well.	
35. In order for me to do something important, I have to have all my doubts worked out.	<b>Reverse</b>

## Appendix B

### The Beck Depression Inventory (BDI-II)

<b>BDI-II</b>	Date: _____
---------------	-------------

Name: \_\_\_\_\_ Marital Status: \_\_\_\_\_ Age: \_\_\_\_\_ Sex: \_\_\_\_\_  
Occupation: \_\_\_\_\_ Education: \_\_\_\_\_

**Instructions:** This questionnaire consists of 21 groups of statements. Please read each group of statements carefully, and then pick out the **one statement** in each group that best describes the way you have been feeling during the **past two weeks, including today**. Circle the number beside the statement you have picked. If several statements in the group seem to apply equally well, circle the highest number for that group. Be sure that you do not choose more than one statement for any group, including Item 16 (Changes in Sleeping Pattern) or Item 18 (Changes in Appetite).

<p><b>1. Sadness</b></p> <p>0 I do not feel sad.</p> <p>1 I feel sad much of the time.</p> <p>2 I am sad all the time.</p> <p>3 I am so sad or unhappy that I can't stand it.</p> <p><b>2. Pessimism</b></p> <p>0 I am not discouraged about my future.</p> <p>1 I feel more discouraged about my future than I used to be.</p> <p>2 I do not expect things to work out for me.</p> <p>3 I feel my future is hopeless and will only get worse.</p> <p><b>3. Past Failure</b></p> <p>0 I do not feel like a failure.</p> <p>1 I have failed more than I should have.</p> <p>2 As I look back, I see a lot of failures.</p> <p>3 I feel I am a total failure as a person.</p> <p><b>4. Loss of Pleasure</b></p> <p>0 I get as much pleasure as I ever did from the things I enjoy.</p> <p>1 I don't enjoy things as much as I used to.</p> <p>2 I get very little pleasure from the things I used to enjoy.</p> <p>3 I can't get any pleasure from the things I used to enjoy.</p> <p><b>5. Guilty Feelings</b></p> <p>0 I don't feel particularly guilty.</p> <p>1 I feel guilty over many things I have done or should have done.</p> <p>2 I feel quite guilty most of the time.</p> <p>3 I feel guilty all of the time.</p>	<p><b>6. Punishment Feelings</b></p> <p>0 I don't feel I am being punished.</p> <p>1 I feel I may be punished.</p> <p>2 I expect to be punished.</p> <p>3 I feel I am being punished.</p> <p><b>7. Self-Dislike</b></p> <p>0 I feel the same about myself as ever.</p> <p>1 I have lost confidence in myself.</p> <p>2 I am disappointed in myself.</p> <p>3 I dislike myself.</p> <p><b>8. Self-Criticalness</b></p> <p>0 I don't criticize or blame myself more than usual.</p> <p>1 I am more critical of myself than I used to be.</p> <p>2 I criticize myself for all of my faults.</p> <p>3 I blame myself for everything bad that happens.</p> <p><b>9. Suicidal Thoughts or Wishes</b></p> <p>0 I don't have any thoughts of killing myself.</p> <p>1 I have thoughts of killing myself, but I would not carry them out.</p> <p>2 I would like to kill myself.</p> <p>3 I would kill myself if I had the chance.</p> <p><b>10. Crying</b></p> <p>0 I don't cry anymore than I used to.</p> <p>1 I cry more than I used to.</p> <p>2 I cry over every little thing.</p> <p>3 I feel like crying, but I can't.</p>
--	--

Subtotal Page 1

Continued on Back

 THE PSYCHOLOGICAL CORPORATION  
Harcourt Brace & Company  
SAN ANTONIO  
Orlando • Boston • New York • Chicago • San Francisco • Atlanta • Dallas  
San Diego • Philadelphia • Austin • Fort Worth • Toronto • London • Sydney

Copyright © 1996 by Aaron T. Beck  
All rights reserved. Printed in the United States of America.

0154018392

## Appendix B (continued)

<p><b>11. Agitation</b></p> <p>0 I am no more restless or wound up than usual.</p> <p>1 I feel more restless or wound up than usual.</p> <p>2 I am so restless or agitated that it's hard to stay still.</p> <p>3 I am so restless or agitated that I have to keep moving or doing something.</p> <p><b>12. Loss of Interest</b></p> <p>0 I have not lost interest in other people or activities.</p> <p>1 I am less interested in other people or things than before.</p> <p>2 I have lost most of my interest in other people or things.</p> <p>3 It's hard to get interested in anything.</p> <p><b>13. Indecisiveness</b></p> <p>0 I make decisions about as well as ever.</p> <p>1 I find it more difficult to make decisions than usual.</p> <p>2 I have much greater difficulty in making decisions than I used to.</p> <p>3 I have trouble making any decisions.</p> <p><b>14. Worthlessness</b></p> <p>0 I do not feel I am worthless.</p> <p>1 I don't consider myself as worthwhile and useful as I used to.</p> <p>2 I feel more worthless as compared to other people.</p> <p>3 I feel utterly worthless.</p> <p><b>15. Loss of Energy</b></p> <p>0 I have as much energy as ever.</p> <p>1 I have less energy than I used to have.</p> <p>2 I don't have enough energy to do very much.</p> <p>3 I don't have enough energy to do anything.</p> <p><b>16. Changes in Sleeping Pattern</b></p> <p>0 I have not experienced any change in my sleeping pattern.</p> <hr/> <p>1a I sleep somewhat more than usual.</p> <hr/> <p>1b I sleep somewhat less than usual.</p> <hr/> <p>2a I sleep a lot more than usual.</p> <hr/> <p>2b I sleep a lot less than usual.</p> <hr/> <p>3a I sleep most of the day.</p> <hr/> <p>3b I wake up 1-2 hours early and can't get back to sleep.</p>	<p><b>17. Irritability</b></p> <p>0 I am no more irritable than usual.</p> <p>1 I am more irritable than usual.</p> <p>2 I am much more irritable than usual.</p> <p>3 I am irritable all the time.</p> <p><b>18. Changes in Appetite</b></p> <p>0 I have not experienced any change in my appetite.</p> <hr/> <p>1a My appetite is somewhat less than usual.</p> <hr/> <p>1b My appetite is somewhat greater than usual.</p> <hr/> <p>2a My appetite is much less than before.</p> <hr/> <p>2b My appetite is much greater than usual.</p> <hr/> <p>3a I have no appetite at all.</p> <hr/> <p>3b I crave food all the time.</p> <p><b>19. Concentration Difficulty</b></p> <p>0 I can concentrate as well as ever.</p> <p>1 I can't concentrate as well as usual.</p> <p>2 It's hard to keep my mind on anything for very long.</p> <p>3 I find I can't concentrate on anything.</p> <p><b>20. Tiredness or Fatigue</b></p> <p>0 I am no more tired or fatigued than usual.</p> <p>1 I get more tired or fatigued more easily than usual.</p> <p>2 I am too tired or fatigued to do a lot of the things I used to do.</p> <p>3 I am too tired or fatigued to do most of the things I used to do.</p> <p><b>21. Loss of Interest in Sex</b></p> <p>0 I have not noticed any recent change in my interest in sex.</p> <p>1 I am less interested in sex than I used to be.</p> <p>2 I am much less interested in sex now.</p> <p>3 I have lost interest in sex completely.</p>
--	---

6789101112 ABCDE

**NOTICE:** This form is printed with both blue and black ink. If your copy does not appear this way, it has been photocopied in violation of copyright laws.

Subtotal Page 2

Subtotal Page 1

Total Score

## Appendix C

### The State Trait Anxiety Inventory (STAI Form Y-2)

#### SELF-EVALUATION QUESTIONNAIRE

STAI Form Y-2

Name \_\_\_\_\_ Date \_\_\_\_\_

#### DIRECTIONS

A number of statements which people have used to describe themselves are given below. Read each statement and then circle the appropriate number to the right of the statement to indicate how you *generally* feel. There are no right or wrong answers. Do not spend too much time on any one statement but give the answer which seems to describe how you generally feel.

ALMOST NEVER  
SOMETIMES  
OFTEN  
ALMOST ALWAYS

- |   |   |   |   |   |
|---|---|---|---|---|
| 21. I feel pleasant .....   | 1 | 2 | 3 | 4 |
| 22. I feel nervous and restless .....   | 1 | 2 | 3 | 4 |
| 23. I feel satisfied with myself .....  | 1 | 2 | 3 | 4 |
| 24. I wish I could be as happy as others seem to be .....   | 1 | 2 | 3 | 4 |
| 25. I feel like a failure .....   | 1 | 2 | 3 | 4 |
| 26. I feel rested .....   | 1 | 2 | 3 | 4 |
| 27. I am "calm, cool, and collected" .....  | 1 | 2 | 3 | 4 |
| 28. I feel that difficulties are piling up so that I cannot overcome them .....                   | 1 | 2 | 3 | 4 |
| 29. I worry too much over something that really doesn't matter .....                              | 1 | 2 | 3 | 4 |
| 30. I am happy .....  | 1 | 2 | 3 | 4 |
| 31. I have disturbing thoughts .....  | 1 | 2 | 3 | 4 |
| 32. I lack self-confidence .....  | 1 | 2 | 3 | 4 |
| 33. I feel secure .....   | 1 | 2 | 3 | 4 |
| 34. I make decisions easily .....   | 1 | 2 | 3 | 4 |
| 35. I feel inadequate .....   | 1 | 2 | 3 | 4 |
| 36. I am content .....  | 1 | 2 | 3 | 4 |
| 37. Some unimportant thought runs through my mind and bothers me .....                            | 1 | 2 | 3 | 4 |
| 38. I take disappointments so keenly that I can't put them out of my mind .....                   | 1 | 2 | 3 | 4 |
| 39. I am a steady person .....  | 1 | 2 | 3 | 4 |
| 40. I get in a state of tension or turmoil as I think over my recent concerns and interests ..... | 1 | 2 | 3 | 4 |

© Copyright 1968, 1977 by Charles D. Spielberger. All rights reserved.  
Published by Mind Garden, Inc., 1690 Woodside Rd, Suite 202, Redwood City, CA 94061

STAI-AD Test Form \\  
www.mindgarden.com

Note: Scores on items 21, 23, 26, 27, 30, 33, 34, 36 and 39 are reversed

## Appendix D

### Mean Ratings of the Forty Self-statements (Twenty Positive and Twenty Negative) used in Experiment 1

Self-Statement	Mean Rating
1* I love life	18.65
2* I know that I am loved	17.79
3* I am happy with who I am	17.77
4* There is so much that I can do with my life	17.48
5* There is so much for me to be happy about	17.41
6* I am part of a beautiful world	17.35
7* When things go wrong I know that I will always have friends	16.92
8* I have no problems that can't be solved	16.73
9* I am proud of myself	16.68
10* I am whole	16.37
11 I can be whatever I want	16.32
12 I can turn my failure into success	16.18
13 Inside I am a good person	15.72
14 I have strength in my beliefs	15.66
15 People like me	15.53
16 I have good qualities	15.40
17 I am lucky	15.39
18 There are things about me that I like	15.13
19 There is always hope for my life	15.13
20 My heart is in the right place	14.44
21 There are things about me I can never say	7.84
22 I have problems people don't know about	7.66
23 I thought I would be better than this	6.92
24 I am not as good as people think	6.48
25 I can't help feeling lonely	5.85
26 No one really understands me	5.10
27 People don't really like me	4.26
28 I am ashamed of myself	3.81
29 Sometimes I feel worthless	3.27
30* Deep down there is something wrong with me	3.25
31* No one will ever love me	3.23
32* Sometimes I wish I wasn't me	3.21
33* I am helpless	3.07
34* I am ugly	2.85
35* I am broken	2.71
36* I make a mess of everything	2.69
37* I am a bad person	2.66
38* I am stupid	2.54
39* My life is pointless	2.39
40* I am a failure	1.45

\* indicates that self-statement was actually employed in the experiment

## Appendix E

### Consent Form Employed in Experiment 1

#### INFORMED CONSENT FORM

**In agreeing to participate in this research I understand the following:**

This research is being conducted by CLAIRE KEOGH, a Ph.D. student at the Department of Psychology, National University of Ireland, Maynooth (NUIM).

- I understand that it is Ms. Keogh's responsibility to adhere to ethical guidelines in her dealings with participants and in the collection and handling of data.
- I also understand that the activities of Ms. Keogh in this regard are supervised by her doctoral supervisor.
- I have been informed as to the general nature of the study and agree voluntarily to participate.
- I understand that some of the questionnaires contain questions that I may find personal and/or distressing. To this end, I understand that: (1) I may refuse to answer any of the questions without penalty; (2) there are no correct answers to the questions and (3) the questionnaires are simply included to assist the experimenter in interpreting the data. I understand that I may be provided with my overall scores on each of these scales (without interpretation). However, I will not be provided with the scores from any other participant.
- I understand that as part of the experimental procedure, I will be exposed to a number of positive and negative self-statements, some of which I may find distressing.
- In accordance with research guidelines, all data will be kept on file at the Psychology Department in NUIM for approximately 5 years. I give my permission for this.
- I have been made aware that all data from the study will be treated confidentially and that participants will not be identified by name at any stage of the data collection, analysis or write-up (code numbers will be used instead). However, in exceptional circumstances, I understand that my data may be disclosed to Ms. Keogh's doctoral supervisor.
- I understand that I will not be given the full rationale behind the study before my participation, for reasons of best practice. However, it has been made clear that at the conclusion of my participation, any questions or concerns I have will be fully addressed.

**Appendix E (continued)**

- I understand that if any psychological problems arise directly from the research, I can contact Ms. Keogh's doctoral supervisor, Dr. Yvonne Barnes-Holmes, free of charge, to discuss these issues with a Chartered Psychologist.
- I understand that I may withdraw from this study at any time, and may withdraw my data at the conclusion of my participation if I still have concern.

I agree to take part in the above research study.

Participant's Name (PRINT):

---

Researcher's Name (PRINT):

---

Participant's Name (Signed):

---

Researcher's Name (PRINT):

---

Date:

---

Date:

---



## Appendix F

### The Acceptance and Action Questionnaire (AAQ-49)

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 can be dangerous. R	1	2	3	4	5	6	7
2. It is normal to sometimes feel unhappy.	1	2	3	4	5	6	7
3. I can do things that are important to me even when I'm feeling unhappy.	1	2	3	4	5	6	7
4. I can move towards important goals, even if I don't feel good about myself.	1	2	3	4	5	6	7
5. My thoughts and feelings get in the way of my success. R	1	2	3	4	5	6	7
6. If I have mean or nasty thoughts, then I am a mean or nasty person. R	1	2	3	4	5	6	7
7. I try to achieve my goals, even if I am uncertain that I can.	1	2	3	4	5	6	7
8. I try hard not to have bad feelings. R	1	2	3	4	5	6	7
9. I work towards things I value, even though at times I feel uncomfortable or uncertain.	1	2	3	4	5	6	7
10. The greater my worries or anxieties become, the more concerned I get for my well-being. R	1	2	3	4	5	6	7
11. I take action on a problem, even when I fear I may get it wrong.	1	2	3	4	5	6	7
12. The way I feel in a situation usually determines the actions that I take. R	1	2	3	4	5	6	7
13. It's OK for me to have thoughts and feelings that I don't like.	1	2	3	4	5	6	7
14. I am not very aware of what occurs around me. R	1	2	3	4	5	6	7
15. I can set a course in my life and stick to it, even if I have doubts.	1	2	3	4	5	6	7
16. Anxiety is bad. R	1	2	3	4	5	6	7

**Appendix F (continued)**

1	2	3	4	5	6	7	
Never True	Very seldom true	Seldom True	Sometimes True	Frequently True	Almost always true	Always True	
17. My painful experiences and memories make it difficult for me to live a life that I would value. R	1	2	3	4	5	6	7
18. I often catch myself daydreaming about things I've done and what I would do differently next time. R	1	2	3	4	5	6	7
19. I am in control of my life.	1	2	3	4	5	6	7
20. If I get bored of a task, I can still complete it.	1	2	3	4	5	6	7
21. Worries get in the way of my success. R	1	2	3	4	5	6	7
22. If I feel uncertain, I can still make a choice and take action.	1	2	3	4	5	6	7
23. If I promised to do something, I'll do it, even if I later don't feel like it.	1	2	3	4	5	6	7
24. I stop taking care of my responsibilities when I feel anxious or uncomfortable. R	1	2	3	4	5	6	7
25. I try hard to avoid feeling anxious or jittery. R	1	2	3	4	5	6	7
26. In order to achieve my goals, I will not avoid people or places that may upset me.	1	2	3	4	5	6	7
27. Having some worries will not prevent me from living a fulfilling life.	1	2	3	4	5	6	7
28. I should not always believe my reactions and judgments.	1	2	3	4	5	6	7
29. It seems like most people are handling their lives better than I am. R	1	2	3	4	5	6	7
30. I need to control my feelings in order to handle my life well. R	1	2	3	4	5	6	7
31. In order for me to do something important, I first have to have all my doubts worked out. R	1	2	3	4	5	6	7
32. My thoughts and feelings do not get in the way of how I want to live my life.	1	2	3	4	5	6	7
33. I can't stand feeling sad or guilty. R	1	2	3	4	5	6	7

**Appendix F (continued)**

1	2	3	4	5	6	7
never true	very seldom true	seldom true	sometimes true	frequently true	almost always true	always true
34. It's OK if I remember something unpleasant.					1 2 3 4 5 6 7	
35. If an unpleasant memory comes into my head, I try to get rid of it.R					1 2 3 4 5 6 7	
36. I try to avoid thoughts and feelings that cause difficulty in my daily life. R					1 2 3 4 5 6 7	
37. I try hard to control the physical reactions that I experience in my body (e.g., heart racing, sweating). R					1 2 3 4 5 6 7	
38. I would rather achieve my goals than avoid unpleasant thoughts and feelings.					1 2 3 4 5 6 7	
39. Emotions cause problems in my life. R					1 2 3 4 5 6 7	
40. I'm afraid of my feelings. R					1 2 3 4 5 6 7	
41. My painful memories prevent me from having a fulfilling life. R					1 2 3 4 5 6 7	
42. When I feel uneasy, I do whatever I can to get rid of those feelings. R					1 2 3 4 5 6 7	
43. I do not have to control my thoughts and feelings to be successful.					1 2 3 4 5 6 7	
44. I don't avoid situations that make me feel jittery.					1 2 3 4 5 6 7	
45. I don't have to get rid of scary or unhappy images that come to my mind.					1 2 3 4 5 6 7	
46. If I notice myself breathing quickly, then something is wrong. R					1 2 3 4 5 6 7	
47. My mind is often on "automatic pilot", not fully involved in what I am doing in the moment. R					1 2 3 4 5 6 7	
48. I worry about not being able to control my worries and feelings. R					1 2 3 4 5 6 7	
49. It's OK to feel sad or anxious.					1 2 3 4 5 6 7	

*Note.* R indicates that scoring on this item is reversed.

## Appendix G

### The Depression, Anxiety and Stress Scales (DASS-21)

<b>DASS21</b>				
Please read each statement and circle a number 0, 1, 2 or 3 which indicates how much the statement applied to you <i>over the past week</i> . There are no right or wrong answers. Do not spend too much time on any statement.				
<i>The rating scale is as follows:</i>				
0 Did not apply to me at all				
1 Applied to me to some degree, or some of the time				
2 Applied to me to a considerable degree, or a good part of time				
3 Applied to me very much, or most of the time				
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 physical exertion (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

*Note.* DASS: Depression = items 3, 5, 10, 13, 16,17 and 21; Anxiety = items 2, 4, 7, 9, 15, 19, 20; and Sress = 1, 6, 8, 11, 12, 14, 18.

## Appendix H

### The Balanced Inventory of Desirable Responding (BIDR)

Using the scale below as a guide, enter a number beside each statement to indicate how much you AGREE with it for you:

1-----2-----3-----4-----5-----6-----7  
**NOT** **SOMEWHAT or** **VERY**  
**TRUE** **SOMETIMES** **TRUE**

<b>1. My first impressions of people usually turn out to be right</b>	
<b>2. It would be hard for me to break any of my bad habits.</b>	
<b>3. I don't care to know what other people really think of me.</b>	
<b>4. I have not always been honest with myself.</b>	
<b>5. I always know why I like things.</b>	
<b>6. When my emotions are aroused, it biases my thinking.</b>	
<b>7. Once I've made up my mind, other people can seldom change my opinion.</b>	
<b>8. I am not a safe driver when I exceed the speed limit.</b>	
<b>9. I am fully in control of my own fate.</b>	
<b>10. It's hard for me to shut off a disturbing thought.</b>	
<b>11. I never regret my decisions.</b>	
<b>12. I sometimes lose out on things because I can't make up my mind soon enough.</b>	
<b>13. The reason I vote is because my vote can make a difference.</b>	
<b>14. My parents were not always fair when they punished me.</b>	
<b>15. I am a completely rational person.</b>	
<b>16. I rarely appreciate criticism.</b>	
<b>17. I am very confident of my judgments.</b>	
<b>18. I have sometimes doubted my ability as a lover.</b>	
<b>19. It's all right with me if some people happen to dislike me.</b>	

**Appendix H (continued)**

1-----2-----3-----4-----5-----6-----7	
NOT	SOMEWHAT or
TRUE	SOMETIMES
	VERY
	TRUE
<b>20. I don't always know the reasons why I do the things I do.</b>	
<b>21. I sometimes tell lies if I have to.</b>	
<b>22. I never cover up my mistakes.</b>	
<b>23. There have been occasions when I have taken advantage of someone.</b>	
<b>24. I never swear.</b>	
<b>25. I sometimes try to get even rather than forgive and forget.</b>	
<b>26. I always obey laws, even if I'm unlikely to get caught.</b>	
<b>27. I have said something bad about a friend behind his or her back.</b>	
<b>28. When I hear people talking privately, I avoid listening.</b>	
<b>29. I have received too much change from a salesperson without telling him or her.</b>	
<b>30. I always declare everything at customs.</b>	
<b>31. When I was young I sometimes stole things.</b>	
<b>32. I have never dropped litter on the street.</b>	
<b>33. I sometimes drive faster than the speed limit.</b>	
<b>34. I never read sexy books or magazines.</b>	
<b>35. I have done things that I don't tell other people about.</b>	
<b>36. I never take things that don't belong to me.</b>	
<b>37. I have taken sick-leave from work or school even though I wasn't really sick.</b>	
<b>38. I have never damaged a library book or store merchandise without reporting it.</b>	
<b>39. I have some pretty awful habits.</b>	
<b>40. I don't gossip about other people's business</b>	

*Note.* Odd numbered items comprise the self-deception scales, while even numbered items comprise the impression management scales.

## Appendix I

### The Rosenberg Self-Esteem Scale (RSES)

Instructions: Below is a list of statements dealing with your general feelings about yourself. If you strongly agree, circle **SA**. If you agree with the statement, circle **A**. If you disagree, circle **D**. If you strongly disagree, circle **SD**.

1.	On the whole, I am satisfied with myself.	SA	A	D	SD
2.*	At times, I think I am no good at all.	SA	A	D	SD
3.	I feel that I have a number of good qualities.	SA	A	D	SD
4.	I am able to do things as well as most other people.	SA	A	D	SD
5.*	I feel I do not have much to be proud of.	SA	A	D	SD
6.*	I certainly feel useless at times.	SA	A	D	SD
7.	I feel that I'm a person of worth, at least on an equal plane with others.	SA	A	D	SD
8.*	I wish I could have more respect for myself.	SA	A	D	SD
9.*	All in all, I am inclined to feel that I am a failure.	SA	A	D	SD
10.	I take a positive attitude toward myself.	SA	A	D	SD

\* indicates that item is negatively reversed.

## Appendix J

### The Acceptance and Action Questionnaire-II (AAQ-II)

#### AAQ-II

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. Its OK if I remember something unpleasant.	1	2	3	4	5	6	7
2. My painful experiences and memories make it difficult for me to live a life that I would value. R	1	2	3	4	5	6	7
3. I'm afraid of my feelings. R	1	2	3	4	5	6	7
4. I worry about not being able to control my worries and feelings. R	1	2	3	4	5	6	7
5. My painful memories prevent me from having a fulfilling life. R	1	2	3	4	5	6	7
6. I am in control of my life.	1	2	3	4	5	6	7
7. Emotions cause problems in my life. R	1	2	3	4	5	6	7
8. It seems like most people are handling their lives better than I am. R	1	2	3	4	5	6	7
9. Worries get in the way of my success. R	1	2	3	4	5	6	7
10. My thoughts and feelings do not get in the way of how I want to live my life.	1	2	3	4	5	6	7

*Note.* R indicates that scoring on this item is reversed.



## Appendix K

### An Example of an IRAP-Related Response Scale for Experiments 3, 4 and 5

Q1. Circle the number on the scales below that most accurately represents **your** feelings towards each of these self-relevant words.

	<b>Extremely Negative</b>				<b>Extremely Positive</b>			
<b>Kind</b>	1	2	3	4	5	6	7	
<b>Confident</b>	1	2	3	4	5	6	7	
<b>Trusting</b>	1	2	3	4	5	6	7	
<b>Honest</b>	1	2	3	4	5	6	7	
<b>Secure</b>	1	2	3	4	5	6	7	
<b>Popular</b>	1	2	3	4	5	6	7	
<b>Selfish</b>	1	2	3	4	5	6	7	
<b>Self-conscious</b>	1	2	3	4	5	6	7	
<b>Jealous</b>	1	2	3	4	5	6	7	
<b>Fake</b>	1	2	3	4	5	6	7	
<b>Insecure</b>	1	2	3	4	5	6	7	
<b>Lonely</b>	1	2	3	4	5	6	7	

## Appendix L

### The Philadelphia Mindfulness Scales (PHLMS)

#### PHLMS©

**Instructions:** Please circle how often you experienced each of the following statements *within the past week*.

1. I am aware of what thoughts are passing through my mind.

1	2	3	4	5
Never	Rarely	Sometimes	Often	Very Often

2. I try to distract myself when I feel unpleasant emotions.

1	2	3	4	5
Never	Rarely	Sometimes	Often	Very Often

3. When talking with other people, I am aware of their facial and body expressions.

1	2	3	4	5
Never	Rarely	Sometimes	Often	Very Often

4. There are aspects of myself I don't want to think about.

1	2	3	4	5
Never	Rarely	Sometimes	Often	Very Often

5. When I shower, I am aware of how the water is running over my body.

1	2	3	4	5
Never	Rarely	Sometimes	Often	Very Often

6. I try to stay busy to keep thoughts or feelings from coming to mind.

1	2	3	4	5
Never	Rarely	Sometimes	Often	Very Often

7. When I am startled, I notice what is going on inside my body.

1	2	3	4	5
Never	Rarely	Sometimes	Often	Very Often

8. I wish I could control my emotions more easily.

1	2	3	4	5
Never	Rarely	Sometimes	Often	Very Often

9. When I walk outside, I am aware of smells or how the air feels against my face.

1	2	3	4	5
Never	Rarely	Sometimes	Often	Very Often

### Appendix L (continued)

10. I tell myself that I shouldn't have certain thoughts.

1	2	3	4	5
Never	Rarely	Sometimes	Often	Very Often

11. When someone asks how I am feeling, I can identify my emotions easily.

1	2	3	4	5
Never	Rarely	Sometimes	Often	Very Often

12. There are things I try not to think about.

1	2	3	4	5
Never	Rarely	Sometimes	Often	Very Often

13. I am aware of thoughts I'm having when my mood changes.

1	2	3	4	5
Never	Rarely	Sometimes	Often	Very Often

14. I tell myself that I shouldn't feel sad.

1	2	3	4	5
Never	Rarely	Sometimes	Often	Very Often

15. I notice changes inside my body, like my heart beating faster or my muscles getting tense.

1	2	3	4	5
Never	Rarely	Sometimes	Often	Very Often

16. If there is something I don't want to think about, I'll try many things to get it out of my mind.

1	2	3	4	5
Never	Rarely	Sometimes	Often	Very Often

17. Whenever my emotions change, I am conscious of them immediately.

1	2	3	4	5
Never	Rarely	Sometimes	Often	Very Often

18. I try to put my problems out of mind.

1	2	3	4	5
Never	Rarely	Sometimes	Often	Very Often

19. When talking with other people, I am aware of the emotions I am experiencing.

1	2	3	4	5
Never	Rarely	Sometimes	Often	Very Often

20. When I have a bad memory, I try to distract myself to make it go away.

1                      2                      3                      4                      5  
Never                  Rarely                  Sometimes                  Often                  Very Often

### Appendix M

#### Some Examples from the IRAP-Related Believability Scale for Experiments 7 and 8

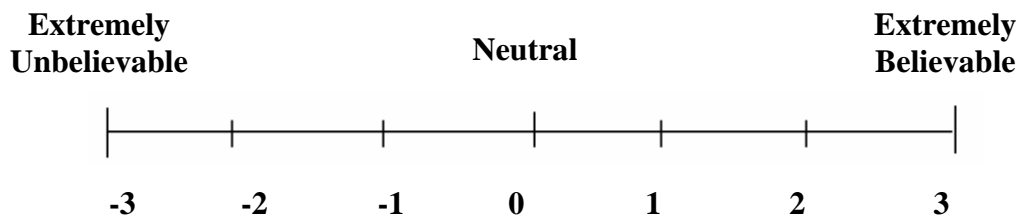
Below you will see a number of rating scales that ask you to rate how you feel about a number of different statements.

Imagine each statement said something about who you are as a person.

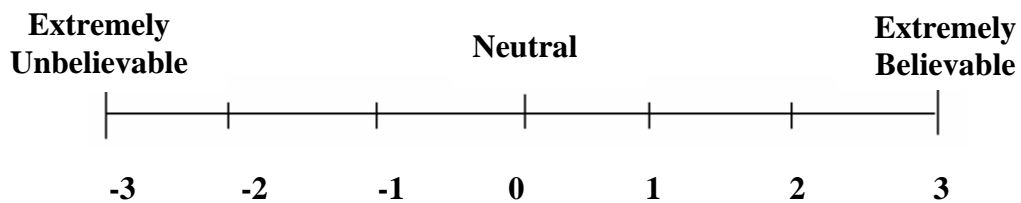
How **BELIEVABLE** do you think each statement is in relation to how you see yourself **right now**? (i.e. How true do you think they are about you **right now**?)

Please take a moment to read each statement to yourself before circling your choice for that statement

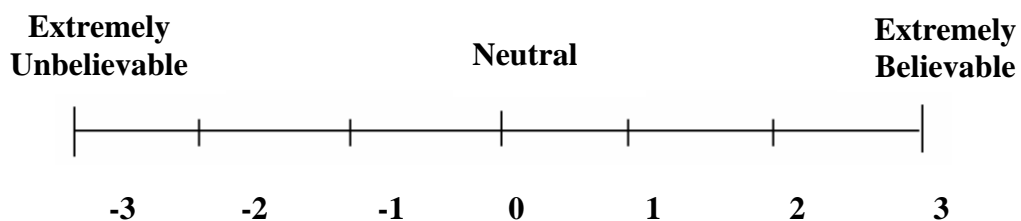
-----  
"I am unpleasant to be around"



"I am afraid to show who I really am"



-----  
"I am jealous of those around me"



## **Appendix N**

### **Experimental Guide to Complete the Online IRAP in Experiment 7**

Many thanks for agreeing to take part in this research.

Before beginning these tasks please make sure that you are on a computer that has software that can properly read Microsoft word files.

Please ensure that you complete this experiment in a quiet room, free from distraction. In addition, please ensure that you complete all aspects of the experiment in one sitting.

Your unique I.D. and password is presented below. This I.D. and password is just for you, and as such must not be used by any other individuals completing this experiment. You will only be required to use the I.D. and password during Step 3 of the experiment (see below).

ID:

PASSWORD:

There are four important steps to completing this experiment:

Step 1: Informed consent

Step 2: Background questionnaire, psychological measures and believability ratings

Step 3: IRAP task

Step 4: Final rating scales

These documents are attached to this email in the order outlined above. Please download each of the documents (for example, to your desktop etc.) and save any changes you make to the documents before returning them to me via email (claire.keogh@gmail.com).

Step 1: Please take a moment to read and sign the attached informed consent form. This consent form must be read carefully and signed before you proceed with the experiment.

Following this, if you are happy to proceed, please complete each of the following experimental steps.

Step 2: This phase requires you to complete a number of brief questions to determine your experience with both ACT and the IRAP. In addition, you will be asked to complete three brief psychological measures, namely the Depression, Anxiety and Stress Scales (DASS21; Lovibond & Lovibond, 1995 or see [www.psy.unsw.edu.au/dass/](http://www.psy.unsw.edu.au/dass/)), the Rosenberg Self-Esteem Questionnaire (RSES;

**Appendix N (continued)**

Rosenberg, 1965) and the The Philadelphia Mindfulness Scales (PHLMS; Cardaciotto, Herbert, Forman, Moitra, & Farrow, in press). Finally, you will be asked to rate the believability of twelve individual statements. Further instructions about completing these measures are provided at the start of the document entitled 'Step 2'.

Step 3: This phase requires you to complete the IRAP computer task. You will need your unique I.D. and password to complete this phase. Detailed instructions as to how to complete the IRAP are provided in the document entitled 'Step 3'. Even if you are familiar with the IRAP, it might be beneficial to have a brief look over the instructions provided here. **This document also provides the link that will take you to the online IRAP website**, where you will complete the IRAP task.

Step 4: Having completed the IRAP computer task, the final step in this experiment requires you to complete a number of brief rating scales. It is important that you complete this section immediately after completing the IRAP task.

Always bear in mind that you are free to discontinue your participation (irrespective of number of steps you have completed) without having to provide an explanation.

Having completed all experimental phases, please return all completed forms (Step 1, Step 2 and Step 4) to me via email at [claire.keogh@gmail.com](mailto:claire.keogh@gmail.com).

If you have any further questions regarding the experiment, or the instructions provided, please do not hesitate to email me at [claire.keogh@gmail.com](mailto:claire.keogh@gmail.com).

Once again, thank you for taking the time to take part in this study,

Warm Regards,  
Claire