The IRAP as a Measure of Implicit Depression and the Role of Psychological Flexibility

Ian Hussey* & Dermot Barnes-Holmes

National University of Ireland Maynooth

*Corresponding Author
Ian Hussey
Department of Psychology
John Hume Building
National University of Ireland Maynooth
Co. Kildare
Ireland
Phone: ++353 (1) 708 6086
Email: Ian.Hussey@nuim.ie

Abstract

A broad implicit measure of depressive emotional reactions was created by mapping the content of the depression scale from the Depression Anxiety and Stress Scale (DASS) on to the Implicit Relational Assessment Procedure (IRAP). Participants were asked to relate pairings of antecedents and emotional reactions that followed the formula 'When X happens…I feel Y.' Groups of participants representing the low and high extremes of normative levels of depressive symptoms completed an IRAP before and after a sad mood induction procedure. At baseline both groups produced a positive emotional response bias on the IRAP. After the sad mood induction, the 'normal' group showed no change, whereas the 'mild/moderate' depression group showed a significant decrease in the positivity of their emotional responses. A similar pattern of differential change was found when groups were created using scores on the AAQ-II. The findings are related to the broader literature on cognitive reactivity and implications for future research are considered.

Keywords: IRAP, implicit cognition, psychological flexibility, mood induction

RUNNING HEAD: Depression, mood induction and the IRAP

Depression is known to be a recurrent disorder that is characterized by negative biases towards the self, the world and the future (Clarke, Beck & Alford, 1999). Remitted depressed individuals who are no longer symptomatic remain to be at increased risk of future depressive episodes, which are triggered by increasingly small environmental stressors (Kendler, Thornton & Gardner, 2000; Mitchell, Parker, Gladstone, Wilhelm & Austin, 2003). Efforts to isolate the vulnerability factors and causal mechanisms involved in the onset, maintenance and relapse of depression have continued to develop over the past thirty years.

Recently, particular attention has been paid to 'cognitive reactivity,' the idea that "sad moods [are] likely to reactivate thinking styles associated with previous sad moods" (Segal, Williams & Teasdale, 2002, p.19). That is, dysfunctional patterns of behavior that emerge within the context of a sad mood state. Reactivity to sad mood state has been shown to predict relapse (see Lau, Segal & Williams, 2004 for review) and, critically, appears to be malleable within the context of therapy. As such, cognitive reactivity appears to be a therapeutically useful construct. Indeed, Mindfulness-Based Cognitive Therapy (MBCT; Segal et al., 2002) was developed specifically to disrupt the link between cognitive reactivity and depressive relapse. The development of so-called mindfulness skills involves learning to hold depressive thoughts and feelings non-judgmentally, recognize their non-literal nature and undermine reactivity to such thoughts. This emphasis on how clients relate to their thoughts and feelings, rather than attempting to change the form or occurrence of such feelings, is representative of a broader trend within contextual cognitive behavioral therapy approaches such as Acceptance and Commitment Therapy (ACT: Hayes, Strosahl, & Wilson, 1999) and Dialectical Behavior Therapy (DBT: Linehan, 1993).

It therefore appears important to develop research tools to detect such small and subtle cognitive reactivity. There has been some success in this regard, in the development of the LEIDS questionnaire (Van der Does, 2002; Williams, Van der Does, Barnhofer, Crane, &

Segal, 2008), which attempts to measure cognitive reactivity in those with a history of depression. However, all such self-report methods are limited by a client's ability to recognize their own reactivity and the role that it has in precipitating relapse. In effect, the utility of self-report measures is known to be limited in that individuals do not have complete introspective access to the causal processes that drive behavior (Nisbett & Wilson, 1977).

Research suggests that the tools used in studying so called 'implicit' attitudes may be used to investigate psychopathological processes that may occur outside of conscious awareness or are susceptible to social influences (De Houwer, 2002; Wiers, Teachmann & De Houwer, 2007). The most frequently used measure of implicit attitudes is the Implicit Association Test (IAT: Greenwald, McGhee & Schwartz, 1998) and its variants (see Roefs, et al., 2011 for review). Instead of requiring individuals to provide self-reports, these measures instead compare the relative ease (i.e. speed in milliseconds) with which individuals can associate certain pairs of stimuli relative to others. For example, individuals who are faster to pair 'Self' with 'Positive' than 'Negative' are said to have high implicit self-esteem (e.g. Gemar, Segal, Sagrati, & Kennedy, 2001). A recent review shows that implicit measures have shown utility in the study of a variety of DSM-IV Axis 1 diagnostic labels, including depression (Roefs et al., 2011). The structure of implicit measures is consistent with the cognitive theory of depression, which stresses the role of unconscious and automatic private events (Clarke et al., 1999).

Critically, results obtained on implicit measures often do not correlate with those obtained using 'explicit' measures such as questionnaires and semi-structured interviews (Roefs et al., 2011). This is the subject of some debate, with some authors questioning the validity of implicit measures (e.g. LeBel & Paunonen, 2011). Others, however, suggest that the real utility of such measures lies in the unique contribution they can make above and beyond traditional measures (e.g. Barnes-Holmes, Barnes-Holmes, Stewart & Boles, 2010).

For example, several studies have found higher levels of implicit self-esteem in remitted depressed individuals relative to healthy controls (De Raedt, Schacht, Franck & De Houwer, 2006; Gemar et al., 2001). In addition, Steinberg, Karpinski and Alloy (2007) found that implicit self-esteem interacted with stressful life events to predict depressive symptomology. This is significant as although the cognitive theory of depression predicts such an interaction (Clarke et al., 1999) it was not found for explicit self-esteem (using the Rosenberg Self-Esteem Scale). Importantly, Nock et al. (2010) found that implicit suicidal cognitions were significantly more predictive of suicide attempts within six months than traditional explicit measures, including questionnaires and clinician assessment. The study of implicit attitudes and cognitions within experimental psychopathology therefore appears to be a productive avenue of research warranting further study.

The current study differs from previous research within the area in its use of an implicit measure, the Implicit Relational Assessment Procedure (IRAP: Barnes-Holmes et al., 2010), which emerged directly from a modern behavioral account of human language and cognition, Relational Frame Theory (RFT; Hayes, Barnes-Holmes & Roche, 2001). According to RFT, the core elements of human cognition are relational responses. That is, the coming to respond to events in certain ways based on their relationships to other events rather than their formal properties. For example, when asked to choose the 'bigger one,' a child may choose a dime over a penny based on its arbitrary 'value' rather than its physical size. This ability to engage in 'arbitrarily applicable' relational responding appears to be key to the emergence of higher cognitive abilities such as planning, thinking and the verbal construction of self (see Hayes et al., 2001 for an in depth account of RFT).

Relational responding, as defined by RFT, provides the basic building blocks for the model of psychopathology and therapy contained within ACT (Hayes et al., 1999). ACT focuses in part on transforming and undermining the functions of problematic verbal

behavior within psychopathology (see Torneke, 2010 for a book length treatment of RFT and ACT). ACT continues to benefit from an ongoing basic research program within the framework of RFT, such as the positive example set by Hooper, Saunders and McHugh (2010) on the derived generalization of thought suppression.

According to RFT all verbal behavior is relational behavior (Hayes, Barnes-Holmes & Roche, 2001). As such, 'explicit' assessment methods such as questionnaires involve relational responding, by definition. However, in standard verbal practice, individuals' verbal reports (i.e. 'extended and elaborated' relational responses: Barnes-Holmes et al., 2010) are under many sources of contextual control. For example, the phrase "How are you?" has a different function within a therapy session than when greeting a coworker. The IRAP was designed to target the relational responses defined in RFT. Responses on explicit measures, such as questionnaires, are similarly affected by contextual factors such as socially desirable responding (see Paulhus, 2002). As an implicit measure, the IRAP is designed to limit contaminating sources of contextual control such as socially desirable responding. In this manner, it is said to specifically target relatively 'brief and immediate' relational responses (Barnes-Holmes et al., 2010).

The IRAP asks participants to relate words or phrases both quickly and accurately, in a similar manner to other response-latency based measures such as the IAT (Greenwald et al., 1998). Unlike questionnaires, which explicitly ask for individuals' self-reports, the IRAP's output is based on the relative speed with which participants can relate a given pair of stimuli in a given way, and that these differences are due to participants' history of relating these stimuli this way. For example, across a large number of trials, participants should be faster to respond that 'Happy' and 'Pleasant' are 'Similar' than they will that 'Happy' and 'Pleasant' are 'Different,' due to these responses being consistent or inconsistent with a participant's history of relating these stimuli in this way. More formally, the IRAP asks participants to

respond quickly and accurately in ways that either coordinate with or do not coordinate with their pre-experimentally established verbal relations. The IRAP requires individuals to respond to large numbers of trials in alternate directions (i.e. consistent or inconsistent with their history of responding; 'Similar' and 'Different' in the above example) before comparing participants' standardized difference scores across directions. The response time differential between consistent and inconsistent trials (defined as the IRAP effect) is assumed to provide an index of the strength of the relational responses being assessed. The key point here is that it is the relative ease (i.e. speed) with which individuals can relate 'Happy' and 'Pleasant' as 'Similar' or 'Different' that is informative of an individual's history of relating these stimuli this way, and the likelihood of their responding this way in the future.

This account, known as the Relational Elaboration and Coherence (REC) model has been presented in detail in a recent article (Barnes-Holmes et al., 2010). As demonstrated above, this account formulates an explanation in terms of behavioral events that may occur either publicly or privately, rather than by appealing to mediating mental constructs such as associations in memory (see Hughes, Barnes-Holmes, & De Houwer, 2011; cf. APE model: Gawronski & Bodenhausen, 2007).

Previous research into implicit cognition within depression has focused predominantly on self-associations, such as the strength of association between self and worthlessness (Franck, De Raedt, Dereu & Van den Abbeele, 2006) or self and depression (Glashouwer & De Jong, 2010), which have been described as measuring facets of implicit self-esteem (De Raedt et al., 2006; Franck et al., 2006). We suggest that it may be more appropriate to employ an implicit measure that targets emotional reactions to events, rather than self-associations, when exploring the mood state dependent effects that define vulnerability to depression (i.e. cognitive reactivity, see Lau et al., 2004). Thus, the IRAP employed in the current study was designed to target both positive and depressive emotional

reactions to positive and depressing events. The stimuli included in the implicit measure were derived from the content and structure of an established and psychometrically sound explicit measure: the Depression, Anxiety Stress Scale (DASS; Lovibond & Lovibond, 1993). In this manner, the current study differs from previous work in which the rationale for the selection and inclusion of specific stimuli was not stated, as is often the case in implicit attitude research (e.g., De Raedt et al., 2006; Franck et al., 2006; Gemar et al., 2001; Glashouwer & De Jong, 2010; Nock et al., 2010). The current study should be seen as a proof of concept study, insofar as it explores the utility of an implicit measure in the detection of cognitive reactivity within a normative population. Future work may thus explore this reactivity within currently depressed and remitted-depressed populations.

Individuals were screened before recruitment using the DASS. Experimental participants were selected from individuals who represented the high and low extremes of normative-levels of depressive symptoms That is, participants were selected from those who displayed very few or no depressive symptoms, and those who displayed high, but non-clinical, levels of depressive symptoms. These are referred to as the 'normal' and 'mild/moderate' depression groups, in reference to the classification system laid out in the DASS manual (Lovibond & Lovibond, 1993). The DASS also afforded us the possibility to check the specificity of any effects found to depressive symptomatology, rather than psychopathology generally (cf. Glashouwer & De Jong, 2010).

The current study follows the design of previous research into cognitive reactivity (e.g. Gemar et al., 2001). Participants were assessed on the IRAP at baseline and again after an experimentally induced sad mood state (as a corollary of a real-life negative event). This was done with a standard Musical and Autobiographical Recall Mood Induction Procedure, which requires participants to listen to a piece of emotive music while recalling an appropriate personal memory. This procedure has been found to effectively create a sad

mood in a majority of individuals (Clarke & Teasdale, 1985; see Scher et al., 2005 for review). The general hypothesis was that individuals characterized as 'normal' and mild/moderate depressive would show differential reactions to the mood induction procedure.

At this stage it is important to note that the concept of depression is not a functional one, but instead relies on syndromal classification (American Psychiatric Association, 2000). In line with the functional approach adopted in using the IRAP and the REC model of implicit cognition, it seemed important to include an explicit measure that, likewise, defines psychopathology functionally rather than topographically. According to the ACT model, the process of psychological flexibility is a core functional component of many psychopathologies, including depression. That is, their willingness to make contact with their undesirable thoughts and feelings while acting in a manner congruent with their values and goals (Bond et al., 2011). Indeed, there is evidence that many forms of mental health are mediated by psychological flexibility (see Bond et al., 2011; Ruiz, 2010, for reviews). Therefore participants also completed a measure of psychological flexibility, the AAQ-II (Bond et al., 2011), which allowed us to explore the relationship between this construct and performance on the IRAP.

Method

Participants

A sample of 76 individuals from the Departmental Volunteer Pool were pre-screened using a DASS questionnaire, and a sample of convenience representing the extremes of normative levels of depressive symptoms was recruited. For ethical reasons, individuals who scored above the clinical cut-off for depressive, anxious or stress symptoms on the DASS (depression score ≥ 21, anxiety score ≥ 15, stress scores ≥ 25, Lovibond & Lovibond, 1993) were not eligible for participation. Thirty undergraduate students completed the experimental phases of the current study. The sample consisted of 10 males and 20 females aged 18 to 22

(M = 20). All participants had a high level of English fluency. No incentives were offered for participation. Participants were recruited into two groups, representing the high and low extremes of the normative range of depressive symptoms as scored on the DASS depression scale: the 'normal' group consisted of individuals who scored within the 'normal' range (score = < 10, $M = 1.9$, $SD = 1.2$), and the 'mild/moderate' depression group consisted of individuals who scored within the 'mild' or 'moderate' ranges (score = 10 ≥ 20, $M = 14.9$, $SD = 4.25$). Participants' DASS scores are presented in Table 1.

Table 1.

Scores Obtained From the DASS Scales Including the Cut-off Scores Used to Divide Groups

	Depression			Anxiety			Stress		
	Cutoff scores	*n*	*M* (*SD*)	Cutoff scores	*n*	*M* (*SD*)	Cutoff scores	*n*	*M* (*SD*)
'Normal'	0-9	15	1.9 (1.2)	0-7	22	3.7 (4.2)	0-14	15	12.1 (6.4)
'Mild/ moderate'	10-20	15	14.9 (4.3)	8-14	8	10.3 (7.7)	15-25	15	17.2 (3.7)

Materials and Apparatus

Implicit Relational Assessment Procedure. The IRAP program (Barnes-Holmes et al., 2010) presented all instructions and stimuli and recorded all responses. The stimuli employed in the IRAP were derived from the DASS model of depressive symptoms (Lovibond & Lovibond, 1993), which employs seven subscales: dysphoria, hopelessness, devaluation of life, self-deprecation, lack of interest/involvement, anhedonia, and inertia. The content of the DASS depression scale was mapped onto the IRAP along these seven characteristics (presented in Table 2). This took the form of positive and negative antecedents

and emotional responses that followed the general formula "When X happens…I feel Y." The technical limitations of the IRAP require that an even number of stimuli be imputed. As such, two sets of stimuli were formed from the dysphoria subscale and one from every other subscale to create eight stimuli categories total. Dysphoria was selected over the other subscales, as it is the most frequently reported initial symptom (Lovibond & Lovibond, 1993).

Table 2.

The Stimuli Employed in the IRAP, Organized by Label and Target Type

Positive Antecedents	Negative Antecedents	Corresponding DASS sub-scale
When things go well	When things go badly	Dysphoria
When things are good	When things are bad	Dysphoria
When I'm successful	When I fail	Hopelessness
When I have direction	When I'm aimless	Devaluation of Life
When people praise me	When people criticize me	Self Depreciation
When I get involved	When I do nothing	Lack of Interest/Involvement
When good things happen	When bad things happen	Anhedonia
When I get things done	When I put things off	Inertia
Positive Responses	Negative Responses	
I fell happy	I feel sad	Dysphoria
I feel cheerful	I feed depressed	Dysphoria
I feel positive	I feel hopeless	Hopelessness
Life feels meaningful	Life feels meaningless	Devaluation of Life
I fell worthwhile	I feel worthless	Self Depreciation
I feel enthusiastic	I feel uninterested	Lack of Interest/Involvement
I enjoy life	I can't enjoy life	Anhedonia
I have initiative	I have no initiative	Inertia
Response Option 1: TRUE	Response Option 2: FALSE	

Depression Anxiety Stress Scale. The DASS-42 (Lovibond & Lovibond, 1993) consists of 42 items across the three self-report scales for depressive, anxious and stress symptoms, with 14 items per scale. Each item consists of a 4 point Likert scale for

participants to indicate the extent to which each statement applied to them in the previous week (e.g., “I felt I wasn't worth much as a person”).

Acceptance and Action Questionnaire II. The 10-item version of the AAQ-II was included as an exploratory measure. The AAQ-II (Bond et al., 2011) contains 10 self-report items measuring individuals’ psychological flexibility. It asks participants to indicate their agreement with the statement in each item (e.g., “It’s OK if I remember something unpleasant”) using a 7-point Likert scale.

Mood induction procedure. The musical and autobiographical recall mood induction procedure is the standard procedure used in the study of mood state dependent effects (see Scher et al., 2005 for review). Two excerpts of classical music were used to first induce a sad mood state and then remove it: Albinoni's “Adagio in G Minor” for sad mood induction and Mozart’s “Divertimento 136” for positive mood induction. Both pieces have been demonstrated to be effective in producing significant changes in explicit mood in the large majority of individuals (Barnes-Holmes, Barnes-Holmes, Smeets & Luciano, 2004; Cahill et al., 2007; Clarke & Teasdale, 1985). The pieces were digitally trimmed to exactly 7 minutes. The excerpts were presented to participants through headphones using an MP3 player at a volume that participants reported to be both comfortable and clearly audible.

Procedure

The experimental sequence of the study was divided into three phases: 1) baseline measurement previous to mood induction (‘pre’), 2) sad mood induction and 3) measurement post-mood induction (‘post’). In addition to the DASS and AAQ, explicit measures of mood were provided to participants, although the details and results of these are not reported in the current article (full data set is available from the authors on request).

Phase 1: baseline measurement. Participants completed the study individually. During instructional phases of the IRAP the experimenter was seated next to the participant,

at all other times he was seated a distance away from the participant. Participants first completed an AAQ-II. They were then introduced to the IRAP program, which began by presenting a set of instructions to participants that described the screen layout, method of response, and format of the IRAP. Specifically, participants were told that they would be presented with examples of situations and associated feelings along the formula of "When X happens…I feel Y," and to respond to the relationships between these two statements as either "TRUE" or "FALSE." Participants were asked to respond both as quickly and accurately as possible. Furthermore, it was explained that participants would sometimes be asked to respond in a way that was consistent with the intuitively consistent relationship between life events and their emotional responses (e.g., When things are good - I feel cheerful - True), and sometimes in a way that was inconsistent with this relationship (e.g., When I fail - I feel worthwhile - True). The researcher emphasized that this was part of the experiment, and that the general rule of responding either consistently or inconsistently would reverse from block to block, starting with consistent. After reporting that they understood the instructions, participants began the IRAP task.

The IRAP consisted of blocks of 32 trials each. On each trial, an antecedent stimulus (e.g., 'When things go well') appeared at the top of the screen, an emotional-response stimulus (e.g., 'I feel happy') appeared in the middle and the two response options ('True' and 'False') appeared at the bottom left and right hand corners (see Figure 1). Participants were required to choose one of the two response options by pressing the 'D' key (for the left option) or 'K' key (for the right option). If a participant chose the correct response option, the screen cleared and the next trial was presented. If the participant chose the incorrect option a red 'X' appeared in the middle of the screen and remained there until the participant emitted the correct response. If a participant failed to respond within 3000ms from the start of a trial, the warning message "Too Slow" appeared in red towards the bottom center of the screen.

The message remained there until the participant emitted a response (either correct or incorrect).

Trials presented to participants belonged one of four different 'trial-types': (1) Positive antecedents with positive responses (2) Positive antecedents with negative responses (3) Negative antecedents with positive responses, and (4) Negative antecedents with negative responses (see Figure 1). The IRAP program counterbalanced the combination of stimuli within each block, insofar as each block presents all 32 possible combinations of the stimuli. This is done under two constraints: identical trial-types were not presented twice in succession, and the left-right positions of the response options switched quasi-randomly from trial to trial, appearing in the same left-right positions across in no more than three successive trials.

Blocks on the IRAP are organized into pairs across which participants must answer first in one direction and then the other. In the current study this required participants to respond to the trial "When I fail…I feel worthless" with "True" on odd-numbered blocks, and "False" on even-numbered blocks. Participants were reminded of this change in direction of relational responding after each block with the onscreen message: "Important: during the next phase the previously correct and wrong 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."

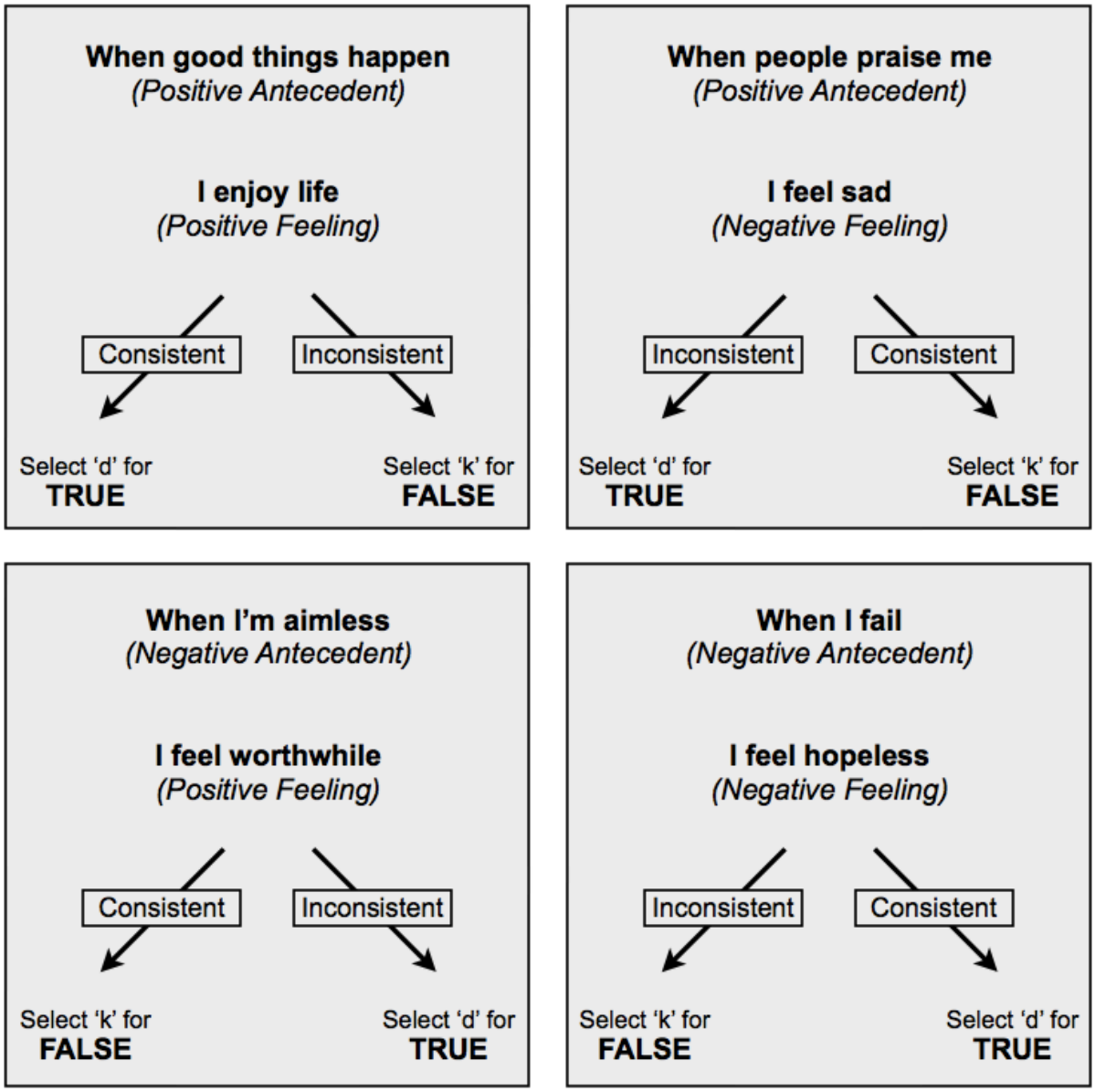

Figure 1. Examples of the four IRAP trial types. Bracketed text, arrows and their labels are included only for illustration and did not appear to participants in the actual IRAP

The IRAP task consisted of between one and three pairs of practice blocks, followed by a fixed set of three pairs of test blocks. Only data from test blocks was included in the analyses. Participants continued from the practice to the test blocks only when their scores on a successive pair of consistent and inconsistent practice blocks reached the criteria for both accuracy (≥ 80% trials correct) and latency (median latency ≤ 3000ms). Feedback on accuracy and latency scores was presented on screen to participants after each block on both

the practice and test blocks. If participants failed to reach accuracy and latency criteria after three pairs of practice blocks they were thanked, debriefed and their data discarded (four participants were excluded from the study on this basis). No performance criteria were used for the test blocks; instead participants' data were excluded if their accuracy fell below 78% (one participant was excluded on this basis), or if their median latency exceeded 2500ms in any given test block (none were excluded on this basis). When all practice and test blocks had been completed, participants were presented with an onscreen message to report to the researcher. Participants were then offered an optional 5-minute break.

Phase 2: sad mood induction. In this phase, a musical mood induction procedure was used establish a sad mood. Participants were provided with the following instructions orally by the experimenter, identical to those used by Barnes-Holmes et al., (2004) and Cahill et al., (2007):

> "You will now be asked to listen to a piece of classical music. This section of music should help you to develop a sad mood. However, music alone cannot create the desired mood, so you should try to think about something that makes you sad. You may find it especially useful to concentrate on sad events that you have personally experienced."

Participants were then provided with a set of headphones through which the sad excerpt (Albinoni's "Adagio in G Minor") was played using the mp3 player.

Phase 3: measurement post-induction. Immediately after the excerpt finished playing participants completed the IRAP for a second time, including both practice and test blocks. All participants completed the second IRAP successfully within the criteria. To ensure that participants left the laboratory in a positive mood context, they received a positive mood induction following the same procedure as the sad mood induction, this time playing Mozart's "Divertimento 136" with an instruction to "Think of something that makes you

happy." When the excerpt had finished, each participant was fully debriefed and thanked for their participation.

Results

The Implicit Relational Assessment Procedure (IRAP)

The primary data produced by the IRAP program are raw latency scores from participants' trials, in milliseconds elapsed between the onset of the trial to the emission of a correct response by the participant. Following standard procedure to control for individual variation (Barnes-Holmes, Waldron, Barnes-Holmes & Stewart, 2009), the response latency data for each participant were transformed into standardized difference scores, or $D\text{-}_{scores}$, using an adaptation of the Greenwald, Nosek, and Banaji (2003) *D*-algorithm (see Cullen & Barnes-Holmes, 2008; Vahey, Barnes-Holmes, Barnes-Holmes & Stewart, 2009). For clarity of interpretation, the scores for the two negative antecedent trial-types were inverted, so that positive $D\text{-}_{scores}$ represent a bias for positive emotional responses and negative scores represent a bias for negative emotional responses.

In effect, a positive emotional response bias on the IRAP indicates that participants, given positive and negative antecedents, are quicker to confirm rather than deny positive emotional reactions; in contrast, a negative emotional response bias indicates that participants are quicker confirm rather than deny negative or depressive emotional reactions. A neutral emotional response bias therefore indicates responding with roughly equal speed when confirming and denying both positive and negative emotional reactions.

Depression groups

A 2 x 4 repeated measures ANOVA was conducted to test for differences between the depression groups at baseline, with group as the between participant variable and IRAP trial-type as the within participant variable. Both the main and interaction effects proved to be non-significant ($p > 0.12$), with both groups demonstrating a similarly positive emotional

response bias. A 2 x 2 x 4 repeated measures ANOVA was then conducted to examine the effect of the sad mood induction procedure. The only significant effect involving the critical pre-post manipulation was an interaction effect with the group variable, F (1, 28) = 8.24, p = 0.01, η_p^2 = 0.2. This interaction is illustrated in Figure 2, upper panel, and indicates that the 'normal' participants showed little change pre to post mood induction, displaying a consistently positive bias. In contrast, the mild/moderate depressive participants showed a substantive reduction post-mood induction.

To explore the interaction effect, participants' overall D-$_{\text{scores}}$ (i.e., response biases) at baseline and post-mood induction were subjected to two independent t-tests, which showed that the two groups did not differ significantly at baseline (p = 0.22), but did following mood induction, t (28) = 2.05, p = 0.05. Two paired t-tests found significant changes in the mild/moderate depression group from baseline to post-mood induction, t (14) = 3.30, p = 0.005; but no significant change in the 'normal' group (p = 0.65).

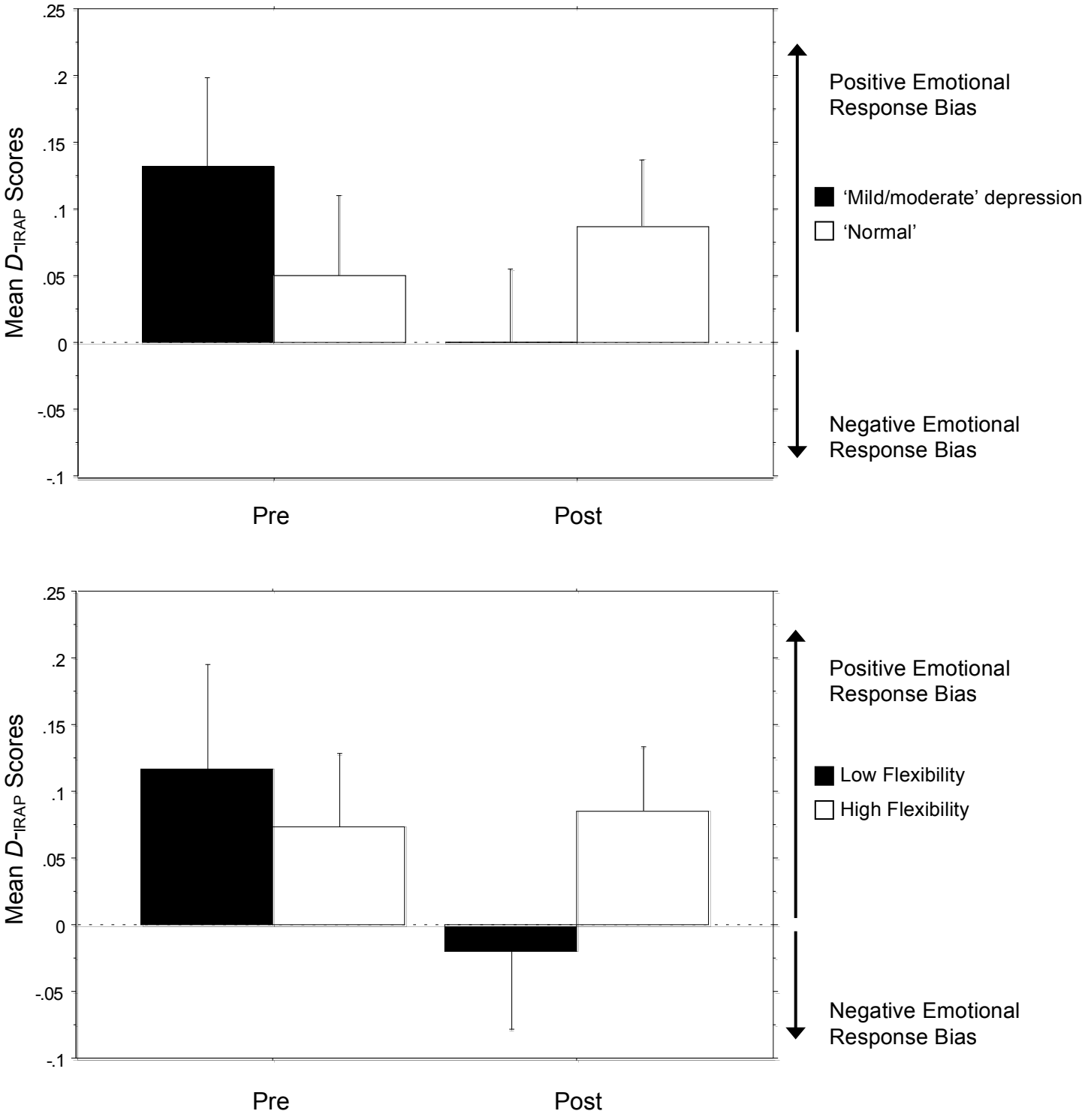

Figure 2. Upper panel: Depression groups' emotional response biases on the IRAP pre and post mood induction. Lower panel: The psychological flexibility groups' emotional response biases on the IRAP pre and post mood induction.

Psychological flexibility groups

The AAQ-II was scored 10-70 (including the reversed scoring of items 2, 3, 4, 5, 7, 8 and 9), with lower scores representing greater levels of psychological flexibility. The AAQ-II was included as an exploratory measure. However, a visual analysis of the distribution of participants' scores on the 10-item AAQ-II showed that they divided cleanly into two groups which were separated by >2 standard errors. As such, participants with AAQ-II scores ≤ 41 were defined post-hoc as the 'high flexibility' group (n = 12, M = 30, SD = 6.27), and those with AAQ-II scores ≥ 48 were defined as the 'low flexibility' group (n = 18, M = 55, SD = 5.04). The scores of these naturally occurring groups are broadly in line with the suggested cut-off scores for healthy versus clinical individuals the AAQ-II (Bond et al., 2011). The overlap between depression and psychological flexibility groups was considerable: all 'normal' individuals were high flexibility, twelve of the fifteen mild/moderate depressed individuals were low flexibility, and the remaining three were high flexibility.

The same analyses that were applied to the depression groups were applied to the psychological flexibility groups. The initial 2 x 4 ANOVA found no differences at baseline (p = 0.54). The 2 x 2 x 4 repeated measures ANOVA again found a single critical interaction effect between pre-post and group, F (1, 28) = 5.76, p = 0.02, η_p^2 = 0.2. The interaction effect is illustrated in Figure 2, lower panel, and shows a similar pattern to that obtained when participants were divided according to their scores on the DASS depression scale. Two independent t-tests confirmed that the two groups did not differ at baseline (p = 0.43), but did post mood induction, t (28) = 2.50, p = 0.02. Paired t-tests confirmed significant change in the low flexibility group from pre to post, t (14) = 2.81, p = 0.02; but not in the high flexibility group (p = 0.77). The analyses demonstrate that the sad mood induction procedure produced a similar change in both mild/moderate depression and low psychological

flexibility individuals, and no significant change in 'normal' and high psychological flexibility individuals.

Specificity of the IRAP to depressive symptoms

To determine if the effects found on the IRAP were attributable to depressive symptoms specifically rather than psychopathology generally, groups were created post-hoc from participants' anxiety and stress scores on the DASS (see Table 1). As with the depression groups, cut-offs for 'normal' and 'mild/moderate' anxiety and stress groups were taken from the DASS manual (Lovibond & Lovibond, 1993). Two 2 x 2 x 4 repeated measures ANOVAs failed to yield any significant interaction effects between group and pre-post mood induction (anxiety, $p > 0.09$; stress, $p > 0.48$). Results on the IRAP therefore appear to be specific to depressive symptoms, as might be expected given that the IRAP's stimuli were mapped from the DASS depression scale.

Discussion

Previously it was noted that the IRAP was designed to target a particular type of the relational responses that, according to Relational Frame Theory, are the basis of cognition: specifically, brief and immediate relational responses. According to the REC model, these are to be contrasted with the 'extended and elaborated' relational responding that is typically reflected in standard verbal practice, such as responses on traditional questionnaires (see Barnes-Holmes et al., 2010 for a more detailed discussion). From this perspective, the results show that the mood induction procedure impacted on the brief relational responses for individuals demonstrating mild/moderate levels of depressive symptoms, but not those demonstrating 'normal' levels of depressive symptoms. Conceptually, this suggests that brief relational responses associated with emotional reactions to life events are relatively sensitive to mood state in individuals characterized as mild/moderate depressive.

Critically, no differences were found between 'normal' and mild/moderate groups

when separated by anxiety or stress scores on the DASS. Effects on the IRAP were therefore attributable to depressive symptoms specifically, rather than psychopathology generally. This is perhaps unsurprising, given that the IRAP's stimuli were derived from the DASS depression scale, and the mood induction is intended to induce a sad mood state. Future research might employ additional measures that are designed to assess the subjective level of distress or discomfort induced by the IRAP itself.

The current study is novel in that it compared groups which were created using scores on a functional measure of psychopathology. The AAQ-II is designed to assess psychological flexibility (Bond et al., 2011). The IRAP itself can be conceptualized as a measure of psychological flexibility, insofar as it appears to be useful in assessing such flexibility within specific response classes that appear to be predictive of future behavior, i.e., brief and immediate relational responses (e.g. Carpenter, Martinez, Vadhan, Barnes-Holmes & Nunes, 2012; Nicholson & Barnes-Holmes, in press). As such, the similar pattern of change found between depression and psychological flexibility groups was, perhaps, unsurprising.

Critically, within the current experiment, individuals who were low in depressive symptoms or high in flexibility showed lower levels of reactivity to external stressors on the IRAP, relative to those who were high in depressive symptoms and low in flexibility. Flexibility here refers not to change per se, but an ability to react to external stressors with attenuated psychopathological responses. More informally, flexibility refers to an individuals ability to take an external stressor "on the chin," whereas inflexibility refers to the tendency to "throw in the towel" when facing psychological challenge. From this perspective, the lack of change seen on the IRAP following sad mood induction indicates high flexibility, whereas change indicates low flexibility. Of course, further research is required to determine to what extent such effects on the IRAP are predictive of psychopathological behaviors in the natural environment, although preliminary research indicates that the IRAP possesses promising

predictive validity (e.g. abstinence within a cocaine treatment program: Carpenter et al., 2012).

These results are broadly consistent with previous research on the impact of mood induction on implicit measures of constructs related to depression (Reifs et al., 2011), insofar as groups demonstrated differential responses to the mood induction. The current study additionally differs from previous work it its use of an implicit measure that allows for analysis of specific relations rather than general associations. That is, it is important to note that the change in the IRAP effect seen in the mild/moderate depressive and high flexibility groups may have been driven by a decrease in positivity, an increase in negativity, or a complex combination of both. As such, the current results suggest that a particular type of brief relational responding is cognitively reactive, at least in the context of the current experiment. Of course, the foregoing remains somewhat speculative, but it does highlight the potential for tightly defined functional analyses of human language and cognition (i.e. RFT and the REC model) to feed into both basic and applied research in human psychopathology, and to create stronger links between behavioral and cognitive approaches. In general, the IRAP therefore appears to have utility as a measure of cognitive reactivity. Future work should seek to parse out and explore the specific relations (i.e. trial-type level effects) implicated in the onset and maintenance of depressive symptoms across a broader range of depressive symptomatology, e.g., those scoring in the ‘severe’ or ‘extremely severe’ range on the DASS.

Finally, it appears important to note that brief and immediate relational responses have been shown to be malleable. Hooper, Villatte, Neofotistou and McHugh (2010) found performance on an IRAP that targeted acceptance or suppression of unwanted thoughts could be altered using a mindful breathing exercise versus a thought-suppression intervention. Given that brief relational responses appear to be highly predictive of clinically relevant

future behaviors, such as onset or relapse (Steinberg et al., 2007), response to treatment (Carpenter et al., 2012), and suicide attempts (Nock et al., 2010), future research should also focus on the malleability of brief and immediate relational responding within therapeutic settings.

Disclosure Statement

None of the authors have a conflict of interest.

References

American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders* (4th ed., text rev.). Washington, DC: Author.

Barnes-Holmes, D., Barnes-Holmes, Y., Stewart, I. & Boles, S. (2010). A sketch of the Implicit Relational Assessment Procedure (IRAP) and the Relational Elaboration and Coherence (REC) model. *The Psychological Record, 60*, 527-542.

Barnes-Holmes, D., Waldron, D., Barnes-Holmes, Y., & Stewart, I. (2009). Testing the validity of the Implicit Relational Assessment Procedure and the Implicit Association Test: Measuring attitudes toward Dublin and country life in Ireland. *The Psychological Record, 59*(3), 389-406.

Barnes-Holmes, Y., Barnes-Holmes, D., Smeets, P. M., & Luciano, C. (2004). The derived transfer of mood functions through equivalence relations. *The Psychological Record, 54,* 95-114.

Bond, F. W., Hayes, S. C., Baer, R. A., Carpenter, K. M., Guenole, N., Orcutt, H. K., Waltz, T., & Zettle, R. D. (2011). Preliminary psychometric properties of the Acceptance and Action Questionnaire II: A revised measure of psychological inflexibility and experiential avoidance. *Behavior Therapy, 42*(4), 676–688.

Cahill, J., Barnes-Holmes, Y., Barnes-Holmes, D., Rodriguez-Valverde, M., Luciano, C., & Smeets, P. M. (2007). The derived transfer and reversal of mood functions through equivalence Relations: II. *The Psychological Record, 57*(3), 373-389.

Carpenter, K. M., Martinez, D., Vadhan, N. L., Barnes-Holmes, D., & Nunes, V. E. (2012). Performance based measures of attentional bias and relational responding are associated with behavioral treatment outcome for cocaine dependence. *The American Journal of Drug and Alcohol Abuse, 38*(2), 146–154. doi:10.3109/00952990.2011.643986

Clarke, D., Beck, A. T. & Alford, B. A. (1999). *Scientific foundations of cognitive theory and therapy of depression,* New York: Wiley.

Clarke, D. & Teasdale, J. (1985). Constraints on the effects of mood on memory. *Journal of Personality and Social Psychology, 48*, 1595-1608.

Cullen, C., & Barnes-Holmes, D. (2008). Implicit pride and prejudice: A heterosexual phenomenon? In T. G. Morrison & M. A. Morrison (Eds.), *Modern Prejudice.* NY: Nova Science.

De Houwer, J. (2002). The Implicit Association Test as a tool for studying dysfunctional associations in psychopathology: Strengths and limitations. *Journal of Behaviour Therapy and Experimental Psychiatry, 33,* 115-133.

De Raedt, R., Schacht, R., Franck, E., & De Houwer, J. (2006). Self-esteem and depression revisited: Implicit positive self-esteem in depressed patients? *Behavior Research and Therapy, 44*, 1017–1028. doi:10.1016/j.brat.2005.08.003

Franck, E., De Raedt, R., Dereu, M., & Van den Abbeele, D. (2006). Implicit and explicit self-esteem in currently depressed individuals with and without suicidal ideation. *Journal of Behavior Therapy and Experimental Psychiatry, 38,* 75–85.

Gawronski, B., & Bodenhausen, G.V. (2007). Unraveling the processes underlying evaluation: Attitudes from the perspective of the APE model. *Social Cognition, 25,* 687-717.

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*(2), 282-289. doi: 10.1037//0021-843X.110.2.282

Glashouwer, K. A., & De Jong, P. J. (2010). Disorder-specific automatic self-associations in depression and anxiety: results of The Netherlands Study of Depression and Anxiety. *Psychological Medicine*, *40*, 1101–1111. doi:10.1017/S0033291709991371

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., & Banaji, M.R. (2003). Understanding and using the Implicit Association Test: An improved scoring algorithm. *Journal of Personality and Social Psychology, 85,* 197-216.

Hayes, S. C. Barnes-Holmes, D., & Roche, B. (Eds.). (2001). *Relational Frame Theory: A Post-Skinnerian account of human language and cognition.* NY: Plenum.

Hayes, S. C., Strosahl, K. D., & Wilson, K. D. (1999). *Acceptance and Commitment Therapy: An experimental approach to behavior change.* NY: Guildford.

Hooper, N., Saunders, J., & McHugh, L. (2010). The derived generalization of thought suppression. *Learning & Behavior, 38*(2), 160-168. doi:10.3758/LB.38.2.160

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

Hughes, S., Barnes-Holmes, D., & De Houwer, J. (2011). The dominance of associative theorising in implicit attitude research: Propositional and behavioral alternatives. *The Psychological Record, 61*, 465–496.

Kendler, K. S., Thornton, L. M., & Gardner, C. O. (2000). Stressful life events and previous episodes in the etiology of major depression in women: An evaluation of the 'kindling' hypothesis. *American Journal of Psychiatry, 157*, 1243–1251.

Lau, M. A., Segal, Z. V., & Williams, J. M. G. (2004). Teasdale's differential activation hypothesis: Implications for mechanisms of depressive relapse and suicidal behaviour. *Behaviour Research and Therapy, 42,* 1001–1017.

LeBel, E. P., & Paunonen, S. V. (2011). Sexy But Often Unreliable: The Impact of Unreliability on the Replicability of Experimental Findings With Implicit Measures. *Personality and Social Psychology Bulletin, 37*(4), 570-583. doi: 10.1177/0146167211400619

Linehan, M. M. (1993). *Cognitive behavioral treatment of borderline personality disorder*. New York: Guilford Press.

Lovibond, S. H. & Lovibond, P. F. (1993). *Manual for the Depression Anxiety Stress Scales (DASS).* Psychology Foundation Monograph.

Mitchell, P. B., Parker, G. B., Gladstone, G. L., Wilhelm, K., & Austin, M. P. V. (2003). Severity of stressful life events in first and subsequent episodes of depression: The relevance of depressive subtype. *Journal of Affective Disorders, 73*, 245–252.

Nicholson, E., & Barnes-Holmes, D. (in press). Developing an Implicit Measure of Disgust Propensity and Disgust Sensitivity: Examining the Role of Implicit Disgust-Propensity and -Sensitivity in Obsessive-Compulsive Tendencies. *Journal of Behavior Therapy and Experimental Psychiatry.* doi:10.1016/j.jbtep.2012.02.001

Nisbett, R. E., & Wilson, T. D. (1977). Telling more than we can know: Verbal reports on mental processes. *Psychological Review, 84,* 231–259. doi:10.1037/0033-295X.84.3.231

Nock, M.K., Park, J.M., Finn, C.T., Deliberto, T.L., Dour, H.J., & Banaji, M.R. (2010). Measuring the Suicidal Mind: Implicit Cognition Predicts Suicidal Behavior. *Psychological Science*, *21*(4), 511-517. doi:10.1177/0956797610364762

Paulhus, D. L. (2002). Socially Desirable Responding: the evolution of a construct, in H. I. Braun, D. N. Jackson, & D. E. Wiley (Eds), *The role of constructs in psychological and educational measurement* (pp. 49-69). Mahwah NJ: Erlbaum.

Roefs, A., Huijding, J., Smulders, F. T. Y., Macleod, C. M., de Jong, P. J., Wiers, R. W. & Jansen, A. T. M. (2011). Implicit measures of association in psychopathology research. *Psychological Bulletin, 137*(1), 149–193. doi:10.1037/a0021729

Ruiz, F. J. (2010). A review of Acceptance and Commitment Therapy (ACT) empirical evidence: correlational, experimental psychopathology, component and outcome studies. *International Journal of Psychology and Psychological Therapy, 10*(1), 125-162.

Scher, C. D., Ingram, R. E. & Segal, Z. V. (2005). Cognitive reactivity and vulnerability: Empirical evaluation of construct activation and cognitive diatheses in unipolar depression. *Clinical Psychology Review, 25,* 487-510. doi:10.1016/j.cpr.2005.01.005

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.

Steinberg, J. A., Karpinski, A., & Alloy, L. B. (2007). The Exploration of Implicit Aspects of Self-esteem in Vulnerability – Stress Models of Depression. *Self and Identity, 6,* 101-117. doi: 10.1080/15298860601118884

Torneke, N. (2010). Learning RFT: An introduction to Relational Frame Theory and its clinical applications. NV: Context Press.

Vahey, N. A., Barnes-Holmes, D, Barnes-Holmes, Y., Stewart, I. (2009). 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, 59*(3), 371-388.

Van Der Does, A. J. W. (2002). Cognitive reactivity to sad mood: structure and validity of a new measure. *Behaviour Research and Therapy, 40,* 105-120.

Wiers, R. W., Teachman, B. A., & De Houwer, J. (2007). Implicit cognitive processes in psychopathology: An introduction, *Journal of Behavior Therapy and Experimental Psychiatry, 38*(2), 95-104. doi:10.1016/j.jbtep.2006.10.002

Williams, J. M. G., Van der Does, A. J. W., Barnhofer, T., Crane, C. & Segal, Z. S. (2008). Cognitive Reactivity, Suicidal Ideation and Future Fluency: Preliminary Investigation of a Differential Activation Theory of Hopelessness/Suicidality. *Cognitive Therapy and Research, 32*, 83–104. doi:10.1007/s10608-006-9105-y