The psychometric properties of the 'Difficulties in Emotional Regulation Scale' in people with schizophrenia and individuals with co-existing depression and substance use disorders

Marianne Ayre

Bachelor of Arts, Graduate Diploma in Education, Graduate Diploma in Counselling, Masters in Education (Educational Psychology)

> Thesis submitted for fulfilment of the award of Doctor of Philosophy (Psychiatry) 2013

STATEMENT OF ORIGINALITY

The thesis contains no material which has been accepted for the award of any other degree or diploma in any university or other tertiary institution and, to the best of my knowledge and belief, contains no material previously published or written by another person, except where due reference has been made in the text. I give consent to the final version of my thesis being made available worldwide when deposited in the University's Digital Repository^{**}, subject to the provisions of the Copyright Act 1968.

I hereby certify that the work embodied in this thesis is the result of original research

Signed _____

Marianne Ayre

Date _____

ACKNOWLEDGEMENTS

This has been a very long journey. I have encountered two car accidents (no! not my fault), which involved great periods of time off with hospital and rehab and ongoing medical attention. So there are many people to thank, especially my supervisor Professor Amanda Baker who stayed with me throughout the long journey. Thank you Amanda for your encouragement and support throughout this process.

I would also like to extend a heart-felt thank you to my second supervisor, Dr Miles Bore. Your insight and guidance gave me the motivation to keep going; I don't think I would have made it otherwise.

A huge thank you to Associate Professor Terry Lewin who is one of the most patient people I know, especially when I made so many blunders losing my database files.

I am grateful to Associate Professor Carmel Loughland who has been a great motivator, and who has always been available with words of encouragement when I doubted myself.

I was also fortunate have the guidance of Dr Frances Kay-Lambkin. Frances, a big thank you for your support and encouragement.

To all the wonderful people who showed an interest in my progress; thank-you. Your gentle nudges and inquiries kept me motivated.

Finally to my wonderful husband Ross Wilson. I love and respect you dearly. I would not have made it without all the meals you cooked and your constant support in everything I dare to dream.

iii

CONTENTS

Statement of Originalityii		
Acknowledgements iii		
Thes	is Overview	1
Chap	ter 1 Introduction	5
1.1	Emotion Regulation	5
1.1.1	Emotional dysregulation	7
1.1.2	Emotional regulation and psychopathology	8
1.1.3	Emotional regulation and psychopathology in individuals with diagnosed	
subst	ance use disorder	9
1.1.4	Emotional regulation and psychopathology in individuals with diagnosed	
Schiz	ophrenia	. 10
1.2	Why do we need a measure?	. 12
1.2.1	What does a measure indicate?	. 13
1.2.2	Multiple measures of constructs related to ER	. 13
1.2.3	Problems with existing scales	. 15
1.3	Studies that report on the DERS	. 17
1.3.1	Limitations of studies using the DERS	. 20
1.3.2	Existing studies on the validity of the DERS	. 22
1.4	Reliability and validity	. 23
1.4.1	Reliability	. 23
1.4.2	Validity	. 24
1.5	Aims of research	. 26
Chap	ter 2 Data Source	. 27
2.1	Overview	. 27
2.2	Introduction	. 27
2.2.1	Summary of DERS measure	. 28
2.2.2	Clinical samples	. 29
2.2.3	Community or control sample (CS)	. 31
2.3	Aims for Chapter Two	. 31
2.4	Methods	. 32
2.4.1	Participants and procedure	. 32
2.4.2	Assessment instruments	. 37
2.4.3	Additional assessment Instruments employed from the SHADE study	. 39
2.4.4	Additional measures from the university student study	. 41
2.4.5	Ethics	. 42
2.4.6	Statistical analysis	. 43
2.5	Results	. 44
2.5.1	Characteristics of the community sample	. 44
2.5.2	Characteristics of clinical samples	. 45
2.6	Discussion	. 52

Chap	ter 3 The psychometric properties of the 'Difficulties in ER scale'	55
3.1	Overview	55
3.2	Introduction	55
3.2.1	Reliability	56
3.2.2	Validity	59
3.3	Aims of Chapter Three	63
3.4	Methods	63
3.4.1	Participants	63
3.4.2	Statistical analysis	64
3.5 3.5.1	<i>Results</i>	64
this s	tudy	64
3.5.2	Reliability of the DERS in an Australian community sample	66
3.5.3	Test-retest reliability	66
3.5.4	Can the DERS detect differences between the control and clinical samples?	68
3.5.5	Factor Analysis	69
3.5.6	Alternative solution on the combined sample	73
3.6	Discussion	80
Chap	oter 4 The Construct Validity of the DERS-R	83
4.1	Overview	83
4.2	Introduction	83
4.2.1	Social problem solving, mood instability, and dysregulation of emotions	84
4.2.2	Big Five from the IPIP (Goldberg, 1992)	86
4.2.3	Rationale for employing two samples in this Chapter	87
4.3	Aim of Chapter Four	87
4.4	Methods	88
4.4.1	Measures	88
4.4.2	Statistical analysis	89
4.5	Results	89
4.5.1	Construct Validity of the DERS-R	89
4.6	Discussion	93
Chap	oter 5 The DERS-R sensitivity to change	95
5.1	Overview	95
5.2	Introduction	95
5.2.1	Alcohol, depression and emotion dysregulation	98
5.3	Aim of Chapter Five	99
5.4	Methods	99
5.4.1	Participants and procedure	99
5.4.2	Measures	99
5.4.3	Statistical analysis 1	101
5.5	Results	102

5.5.1	DERS-R, BDI-II and OTI alcohol mean scores over time of treatment 102		
5.5.2	Trend ana	lysis	104
5.5.3	Random e	ffects model	105
5.6	Discussion		105
Chap	Chapter 6 Discussion		
6.1	The DERS.		109
6.1.1	Summary	of the factor structure of DERS in this thesis	109
6.1.2	Summary	of construct, convergent, and predictive validity of the DERS-	R 110
6.1.3	6.1.3 Limitations and future directions1		
6.1.4	1.4 Strengths of this research		
6.1.5	.1.5 How the DERS-R can be used clinically 1		
6.1.6	Summary.		115
Refe	References		
Appe	endix A-1	DERS – 36 item questionnaire	130
Appe	endix A-2	Information sheet for Hunter Medical Research Institute	
volu	nteers		140
Appendix A-3 and Allied Disor		Information sheet for the Neuroscience Institute of Schize	nhrenia
Appe and A	Allied Diso	rders group	
Appe and A Appe	Allied Diso endix B-1 T	rders group ime 1: Initial information sheet and demographic information group Time 2: Information cover sheet	
Appe and A Appe the co	Allied Diso Allied Diso endix B-1 T ommunity	rders group ime 1: Initial information sheet and demographic information group Time 2: Information cover sheet	on for

TABLE OF TABLES

Table 1-1 Examples of instruments used to report emotional regulation/dysregulation		
Table 1-2 An example of studies using the DERS with number of participants drawn from a clinical population		
Table 2-1 Questionnaires administered to each sample, summary		
Table 2-2 Demographic characteristics of the community sample (CS)		
Table 2-3 Demographic characteristics of the schizophrenia sample		
Table 2-4 Demographic characteristics of CDSA sample 47		
Table 2-5 Depression score with substance use 48		
Table 2-6 Depression and substance severity		
Table 2-7 Demographic characteristics for the combined community and clinical samples 51		
Table 2-8 Demographic characteristics for the university student sample,		
Table 3-1 Internal consistency reported in published articles 58		

Table 3-2 One sample t-test of overall DERS published mean scores and community sample in this thesis. 6	65
Table 3-3 Internal consistency reliability analyses for DERS subscales on CS (n=198) at baseline	56
Table 3-4 Correlations: CS at time 1 (baseline) with time 2 (retest)	67
Table 3-5 Independent t-test comparing control sample with clinical samples	58
Table 3-6 Internal consistency reliability analyses for DERS subscales in combined sample (n=555)	59
Table 3-7 Comparison of six component solution 7	72
Table 3-8 Component analysis (minus items 24, 33)	74
Table 3-9 Item Correlation on the DERS	77
Table 3-10 Final component analysis 7	78
Table 3-11 Correlations between the four subscales	79
Table 4-1 Correlations between scales and factors in combined clinical/non-clinical sample	<i>•</i>
Table 4-2 Correlations between scales and factors in student sample	<i>•</i> 0
Table 4-3 Correlations between the DERS-R and SPSI-R in a combined clinical/non- clinical sample	92
Table 4-4 Correlations between the DERS-R and Big 5 Personality Traits in sample 2.9) 3
Table 5-1 Mean scores, SD, orthogonal polynomial contrasts 10)3
Table 5-2 Correlations between DERS-R and BDI-II)4
Table 5-3 : Correlations indicating change in DERS-R and BDI-II)4

TABLE OF FIGURES

Figure 2:1 Lifetime substance use	. 50
Figure 3:1 Scatter plot of time one and time two	. 67
Figure 3:2 Scree plot using combined data sample	. 70

THESIS OVERVIEW

Difficulties in emotional regulation are considered a core component in over half of the DSM-IV Axis I disorders, and all of the Axis II disorders (Linehan, 1993; Gross & Levenson, 1997; Briere & Gill, 1998; Gratz, 2003). Despite this there are few instruments that measure difficulties in emotional regulation. This has meant that many clinicians have had to use multiple measures that contain closely-related constructs in an attempt to capture the degree of difficulties in emotional regulation experienced by an individual.

One measure that reported sound reliability and validity is the 'Difficulties in Emotional Regulation Scale' (DERS) (Gratz & Roemer, 2004). Since the publication of the DERS, numerous studies have included the measure in research. However the majority of these studies have tended to employ student populations, or report on a small sample size. The few studies that have reported on the psychometric properties of the DERS have indicated problems with one or more of the subscales (Tull & Roemer, 2007; Weinberg & Klonsky, 2009; Bardeen, Fergus, & Orcutt, 2012). This indicates that a re-examination of the psychometric properties of the DERS is needed. It is generally recommended that a re-assessment of the psychometric properties of an existing measure be conducted in different samples under different circumstances and across cultures, to ensure that the measure maintains its robustness and generalisability (Floyd & Widaman, 1995; Haynes, Richard, & Kubany, 1995).

This thesis re-examined the psychometric properties of the DERS within three chapters (Chapters 3 to 5). It aimed to address two of the major limitations reported in the development of the DERS previously, namely, the use of a student sample and, secondly the retest on a small sample size (Gratz et al., 2004). To address these issues this thesis included clinical samples, and second, recruited a large sample for the test-retest reliability. Chapter 1 provided an overview of the literature, whilst Chapter 2 reports the DERS scores across five samples; a community or control sample (CS, n= 198), a university student sample (n=264), and two clinical samples, one with a diagnosis of schizophrenia (n=129) and the second with a co-existing diagnosis of depression and substance abuse (n= 231). The fifth sample was formed by combining the data from the CS, and the clinical groups (n=558). The rationale for combining the

samples was to reflect a sample that is representative of the general population and enabled this study to determine the most appropriate composition of the DERS subscales. An analysis of the total DERS mean scores revealed that the mean in the clinical samples was significantly higher than that of either the CS in this study and the mean score in the university sample reported by Gratz and Roemer (2004). This difference in mean scores indicated that the DERS might be able to detect differences between clinical and community samples and this was examined in Chapter 3.

In study one t-tests were used to compare the DERS total mean score of the CS (n=167) with total mean scores published in other studies, to determine if data from the CS reflected a different population group. Results indicated a statistically significant difference between all the published mean scores and that of the CS, with the CS sample reporting fewer difficulties in emotional regulation. A test of internal consistency of the DERS found the DERS to have high internal reliability and test-retest reliability indicated a minimal degree of change between scores from time one (α =.94) to time two (α =.96). Factor analysis was examined using a diverse sample (n=557) but contrary to expectations, the analysis did not replicate the original six factors of the DERS reported in literature previously. Rather, a principal component analysis with an oblimin rotation found a four factor solution was optimal. Item redundancy was explored and a shortened version of the DERS (DERS-Revised; DERS-R) to be used in an Australian population, was suggested.

A comparison of the DERS-R with existing instruments that are known to reflect a relationship with the construct of the DERS-R will provide further evidence of the validation of the instrument. Chapter 4 examined the construct validity of the DERS-R in two samples. The first sample consisted of clinical and non-clinical participants (n=260) and the second sample consist of university students (n=264). T-tests indicated a significant difference between the samples' mean scores, with the university students reporting greater difficulties in emotional regulation compared to those individuals in the combined sample. To provide evidence for the construct validity of the DERS-R it was expected that the two samples, would report scores on the additional instruments in the same direction.

The Kessler 10 (K10) (Kessler et al., 2002) is a scale of psychological distress, depression and anxiety. It is expected to positively relate to high scores on the DERS-R. Correlations were conducted and both samples revealed a positive correlation in the expected direction between the DERS-R and K10 scores, indicating individuals who experienced greater difficulties in emotional regulation also experienced greater psychological distress. Further examination of the final four factors in both samples showed a statistically significant moderate to high correlation on all four factors of the DERS-R in the expected direction, and all four factors had statistically significant low to moderate correlations with each other in the expected direction.

The 'Social Problem Solving Inventory-Revised' (SPSI-R) (D'Zurilla, Nezu, & Maydeu-Olivares, 1998) was also administered to the combined population to further examine construct validity. SPSI-R reflects difficulties in problem solving in everyday life situations. Problem solving is viewed as an important coping strategy and has been linked with the levels of personal stress or distress a person is experiencing. The scale has two orientations: 'positive' and 'negative', with negative problem solving linked with greater distress. This study hypothesised that a negative relationship would exist between the SPSI-R 'positive orientation' and psychological distress, as measured by K10, and that the DERS-R would be able to reflect this. As predicted, correlation analysis revealed a significant moderate negative relationship between the DERS-R score and 'positive orientation' indicating that poor problem solving was associated with greater difficulties in emotional regulation. Analysis also detected a significant and strong relationship between SPSI-R 'negative orientation' and the DERS-R indicated that greater difficulty in emotional regulation was associated with greater inhibitive cognitive-emotional style, providing further evidence of construct validity.

The student sample was also administered the 'Big 5 Personality Traits Scale' (Goldberg, 1992). On this scale, the trait referred to as 'neuroticism' reflects an individual's tendency to worry, feel insecure, feel nervous and be highly strung. It was hypnotised that someone who experienced greater difficulties in emotional regulation (reflected in the DERS-R) would score high on the 'neuroticism' trait. Analysis confirmed this hypothesis, with a significant high positive relationship on the

neuroticism trait associated with high scores on the DERS-R. The results presented in Chapter 4 provided evidence that the DERS-R has sound construct validity.

Chapter 5 examined whether the DERS-R will be effective in a clinical setting. To do this was necessary to determine if the instrument can detect change, post receiving an intervention. To examine this, a clinical sample who were diagnosed with coexisting depression and alcohol abuse (CDSA, n=103) were recruited, and administered the DERS-R and the Beck Depression Inventory-II (BDI-II) at baseline, 3, 6, and 12 month follow-up. Participants received a 12 week intervention that was designed to reduce alcohol consumption as opposed to specifically focussing on reducing difficulties in emotional regulation. Positive correlations between the DERS-R and the BDI-II were found at each time-period indicating that as individuals experienced greater difficulties in emotional regulation, they also reported greater depression scores.

A random effects model examined the association between BDI-II scores and DERS-R scores at each time period and indicated a significant high association suggesting that each one unit increase in the BDI-II is associated with a 0.89 unit increase in the DERS-R. Finally an examination of the association between the DERS-R and alcohol at each time period revealed a highly significant association, with each one unit increase in alcohol associated with an increase in the DERS-R. This is further evidence of convergent validity. Most importantly, it suggests the DERS-R is able to detect change that occurred as a result of treatment, making it a reliable and effective instrument to use in a clinical population.

The findings from this research have important implications for the measurement of emotional regulation in both the clinical treatment and research setting. The DERS-R was shown to have sound psychometric properties when used in an Australian population. Two major strengths in the DERS-R are that it is slightly briefer that the original DERS, with the DER-R having 29 items compared to the original 36 items. This makes the DERS-R slightly more attractive to clinicians who are required to score the data, and to the clients who are required to complete it, due to a reduction in the time required. Second, the DER-R ability to detect change over-time and post-intervention indicates that it can also be used as an outcome measure.

CHAPTER 1 INTRODUCTION

1.1 Emotion Regulation

Emotional regulation (ER) or affect regulation generally refers to the implicit and explicit efforts that one makes to prolong the positive effects of mood and feeling states while reducing negative emotions. ER is recognised empirically and clinically as having a vital role in the development and maintenance of good mental health. The focus on ER has increased in the past two decades with the emerging field of developmental psychology and attachment theory. Over that time interest in ER has also flourished in adult research especially in mental health disorders.

ER involves a number of diverse systems that require management and organisation including internal (i.e., neurophysiological, cognitive and subjective evaluations), behavioural components (i.e., facial and behavioural actions), and external social components (i.e. cultural values, social contextual significance, personal motivational/goals) (Zeman & Shipman, 1998). There are numerous influences on a child's ability to regulate emotions including previous interactions with the social environment as well as the developmental status and organisation of their emotional systems (Cole, Michel, & O'Donnell, 1994). These influences might enable the individual to either suppress or inhibit emotional arousal or maintain the emotional state. This ability to control one's emotions is viewed as a necessary part of everyday life, especially if one is to be socially accepted (Thompson, 1990). One view espoused by Thompson (1990) places emphasis on the caregivers of a child in helping the child to manage emotional experiences. Being accepted by others is generally thought to be a major index whereby an infant learns to self-regulate. In order to meet the family's conventions, an infant quickly learns to put their tears on hold in order to receive praise, or stop a temper tantrum in order to receive a chocolate. In this way they learn to modulate, tolerate and endure experiences of negative affect. A number of authors support this view (Brazelton & Yogman, 1986; Linehan, 1994) suggesting that successful regulation of distress can foster a sense of mastery. As the child becomes more skilled at regulating emotions they become more adept at social interactions. ER mastery has been associated with an increase in a child's quality of social functioning, cognitive performance of tasks involving delay, habituation, or pursuit of long-term goals, and management of stressful experiences (Frijda, 1988).

Maladaptive regulation is thought to emerge from the invalidation of a child's' experience of emotions, or the experiencing of a trauma or loss in the developmental years of life, and/or having poor external support (Kopp, 1989). If children learn ER in ways that are ineffective they are at risk of subsequent adaptational failures in their social environment (Zeman et al., 1998). Poor ER has been implicated in most forms of childhood psychopathology (Bradley, 2000) and if it is not addressed early, a child will continue to use ineffective ways to regulate emotions into adulthood.

Studies among adults generally describe ER as an individual's ability to influence which emotions they have, when they have them, and how they experience and express them (Gross, 1998b). Functionalist theorists view ER as the individual's ability to flexibly respond to the demands of the environment (Cole et al., 1994; Thompson, 1994). According to functional theorists, ongoing interactions between individuals and the environment are required with the most pertinent interactions occurring in social situations with others. Cole (1994) defines ER as 'the ability to respond to the ongoing demands of experience with the range of emotions in a manner that is socially tolerable and sufficiently flexible to permit spontaneous reactions as well as the ability to delay spontaneous reactions as needed' (p76). Currently most theorists hold to either a one-factor or two-factor theory of ER. Putman and colleagues (Putnam, 2005) state that while there are minor differences between Thompson and Gross' outline of ER, both researchers endorse a 'two-factor' theory. The two-factor theory holds that when intense emotions are experienced there is an attempt by the individual to reduce the emotions by using different techniques or strategies. Other researchers suggest that emotions and ER are not separate or that emotions do not actually precede ER, rather they view emotion and ER as a conjoined process that reflects the attempts by the person to adapt to the situation. This is referred to as onefactor theory, whereby regulation occurs whilst the emotion is activated and can be evident before an emotion is manifested (Putnam, 2005). Although one-factor theorists acknowledge the process outlined in the two-factor theory of ER, they claim the two step process is less frequent than the one step process where emotion and ER occur

simultaneously (Campos, Frankel, & Camras, 2004). While these theories can assist in our understanding of the emotion and ER process, both theories acknowledge adaptation in response to a situation as a component to either the emotion, or the regulation of emotions.

Individuals can be seen to adapt or regulate their emotions in response to their personal experience of discomfort (internal) or the environment itself (external); Examples include avoiding conflict or anger with another, suppressing emotions that are overwhelming, suppressing laughter at a funeral or exam, or following polite norms and demonstrating happiness when not really happy. A study by Gross (1998a) found nine out of ten undergraduate students adapted or regulated their emotions daily.

This thesis will utilize the definition by Gratz and Roemer (2004) who conceptualise ER as involving: (a) awareness and understanding of emotions; (b) acceptance of emotions; (c) ability to control impulsive behaviours and behave in accordance with desired goals when experiencing negative emotions; and (d) ability to use situational-appropriate ER strategies flexibly to modulate emotional responses as desired in order to meet individual goals and situational demands. Emotional dysregulation (ED) tends to occur if one or more of these abilities are absent.

1.1.1 Emotional dysregulation

ED underlies a variety of psychological difficulties and has been linked to difficulties in an individual's functioning across important domains of their lives including work and interpersonal relationships (Gross & Levenson, 1997; Lopes, Salovey, Cote, Beers, & Petty, 2005). ED has also been further implicated in over half of the DSM-IV Axis I disorders and in all of the Axis II disorders (Linehan, 1993; Gross & Levenson, 1997; Briere et al., 1998; Bushman, Baumeister, & Phillips, 2001; Jukupcak, Lisak, & Roemer, 2002; Gratz, 2003).

Definitions of ED are diverse and initially emerged from studies of child development. They include references to interference in the processing of information and events (Dodge, 1989; Calkins, 1994), difficulties with the flexible integration of emotion and other processes (Rieder & Cicchetti, 1989), and poor control over affective experience and expression (Kopp, 1989; Izard, 2002). In general, ED has been attributed to a failure to meet the developmental tasks of emotional development (Dodge, 1989; Kopp, 1989; Izard, 2002). One of the most significant aspects of this is generally thought to be early socio-emotional deprivation (Izard, 2002).

Linehan (1993) states ED is the combination of an emotional response system that is over-sensitive and over-reactive with an inability to modulate the resulting strong emotions and actions associated with them. Generally, all individuals have experienced a time in their life when they have become emotionally dysregulated, it is only when it becomes a problem to the individual or in social interaction with others that it might be considered necessary to provide therapeutic treatments. Examples of everyday experiences of ED include a burst of anger that can turn into a rage or a prolonged sadness that becomes depression. Whilst temper tantrums of a two year old are acceptable, if seen in a six to ten year old, or in an adult, this might be considered problematic (Gross, 2002).

When people are emotionally dysregulated they often report feeling 'out of control' and may say and do things that they would not normally do. The experience of emotion may become overwhelming so they try to regain control or regulate it in a number of ways. Examples of this are to inhibit or suppress emotion(s) through drugs, alcohol or gambling, or via more 'impulsive' behaviours, such as self-harm, shop-lifting, fast driving, extreme sports and so forth (Gross, 2002). Eventually the behaviours used in trying to regain control become out of control and can also lead to experiencing difficulties in everyday life, for example, an inability to hold down a job, studies, or a stable relationship and it is at this point the behaviours can be viewed as problematic.

1.1.2 Emotional regulation and psychopathology in Adults

Work highlighting the association between ED in clinical disorders has been carried out primarily by Linehan (1993), who developed a therapeutic program referred to as Dialectical Behaviour Therapy, for sufferers of borderline personality disorder. Characteristics of borderline personality disorder include recurrent suicidal behaviour and self-harm, impulsivity which comprises of high risk-taking behaviours, affective instability, chronic feelings of emptiness, intense anger and stress-related paranoia (A.P.A., 2000). Prior to Linehan's work, it was generally viewed that individuals with a diagnosis of borderline personality disorder were not only difficult to treat but any form of psychotherapy had little impact on reducing the client's symptomatology. Linehan (1993) proposed that ED is central to people with borderline personality disorder, and redefined suicide attempt and self-harm as a form of ED. This has changed the way clinicians work with borderline personality disorder clients, by focussing on both ED and self-harm as opposed to only self-harm, and by using the treatment model of 'Dialectical Behaviour Therapy'. As a result, treatment has led to better clinical outcomes, for example, one study compared the effectiveness of 'Dialectical Behaviour Therapy' with treatment as usual (TAU) for patients with borderline personality disorder and examined the impact of baseline severity on effectiveness. Fifty-eight women with borderline personality disorder were randomly assigned to either 12 months of 'Dialectical Behaviour Therapy' or TAU. Findings showed that 'Dialectical Behaviour Therapy' was superior to TAU in reducing highrisk behaviours in patients with borderline personality disorder (Verheul et al., 2003). Another randomised controlled study recruited 20 women veterans who met criteria for borderline personality disorder. They were assigned to either 'Dialectical Behaviour Therapy' or TAU for six months. Compared with patients in TAU, those in 'Dialectical Behaviour Therapy' reported significantly greater decrease in suicidal ideation, hopelessness, depression, and anger expression (Koons et al., 2001).

Two other major disorders that share similar ED behaviours and have a high risk of suicide when compared to borderline personality disorder are co-existing depression and substance use disorders (Kessler et al., 1996; Teesson, Lynskey, & Dengenhardt, 2000; Ries, 2006) and schizophrenia (Allebeck, 1989; Felker, Yazel, & Short, 1996; Gut-Fayand, Dervaux, & Olie, 2001; Hunt et al., 2006). These are considered below.

1.1.3 Emotional regulation and psychopathology in individuals diagnosed with depression and substance use disorder

Research has consistently shown that individuals with co-existing major depression and substance use disorders experience more negative outcomes than their counterparts without co-existing disorders (Regier et al., 1990; Sullivan, Fiellin, & O'Connor, 2005; Berking & Wupperman, 2012). It was estimated that 6.2% of respondents in the Australian National Survey of Mental Health and Well Being (NSMHWB, 2009) met the criteria for affective disorder with depression being most prevalent (4.1%), and another 5.1% met the criteria for substance use disorder. Of the 183,900 people who misused drugs nearly every day in the 12 months prior, almost half had a substance use disorder and 31% had an affective disorder (Australian Bureau of Statistics, 2008). Mood disorders such as major depression are characterised by severe affect or emotion related lability, decreased interest in activities, low self-esteem, poor concentration and/or decision-making and feelings of hopelessness (A.P.A., 2000).

There is a strong link between depression and alcohol or other drug (AOD) use problems (Rae, Joyce, Luty, & Mulder, 2002; Kay-Lambkin, Baker, Lewin, & Carr, 2009; Cheetham, Allen, Yucel, & Lubman, 2010). Numerous studies have reported that a major reason people misuse alcohol is to regulate their emotional experiences (Carpenter & Hasin, 1998; Swendsen et al., 2000). Carpenter (1998) reports that the use of alcohol often relieves distress caused by stressful situations, also that feelings of stress often triggers alcohol use. Khantzian (1990) also links poor ER with substance use, suggesting that substance use is a means of dealing with painful affective states for individuals whose ability for self-regulation of emotion is limited. While the consumption of AOD may start off as a social activity for many, those who have a tendency to avoid, blunt or otherwise control negative affective experiences tend to increase their use of AOD. This has been widely documented in numerous studies (Chaney, Roszell, & Cummings, 1982; Cooper, Frone, & Russell, 1995; Stewart & Kushner, 2001).

1.1.4 Emotional regulation and psychopathology in individuals diagnosed with schizophrenia

Schizophrenia is a heterogeneous disorder, with a complex symptom formation and manifestation, for which no atiology is currently known. Bleuler (1950) has described schizophrenia as essentially a splitting of thoughts (cognition) from feelings (emotion) and a "flattening of affect". Clinical and empirical observations provide evidence that a number of emotion identification, expression and behaviour impairments are associated with schizophrenia, including impairments in face processing (Feinberg, Rifkin, Schaffer, & Walker, 1986; Loughland, Carr, & Lewin, 2001; Schneider et al., 2006), prosodic emotion (Edwards, Pattison, Jackson, & Wales, 2001), and emotional experience (Flack, Laird, & Cavallaro, 1999). These impairments are thought to contribute to particular symptom formation in schizophrenia (i.e., hallucinations, delusions) and to impact negatively on the social outcomes of people with the disorder (Aghevli, Blanchard, & Horan, 2003).

Individuals with schizophrenia are thought to have, on average, a lifespan that is up to 18 years shorter than individuals with no mental illness. There is some evidence to suggest increased rates of mortality due to natural causes are around 2.5 times higher for people with schizophrenia than the general population, particularly for cardiovascular, coronary, digestive, endocrine, infectious, genitourinary, neoplastic, nervous, and respiratory diseases (Auquier, Lancon, Rouillon, Lader, & Holmes, 2006; Mitchell & Lord, 2010). Death due to unnatural causes, in particular suicide, is around 12 times higher than in the general population. Patients with schizophrenia represent 25-30% of all psychiatric patients who commit suicide and have an 8-15% risk of dying of a suicide-related cause (Heila et al., 1997; Saha, Chant, & McGrath, 2007). One major risk factor for suicide among patients with schizophrenia is co-existing substance abuse (Regier et al., 1990; Swofford, Kasckow, Scheller-Gilkey, & Inderbitzin, 1996; Hunt et al., 2006), another is depression (Heila et al., 1997; Fenton, 2000).

It has been estimated that 33.7% of people with schizophrenia experience alcohol dependence (Salloum, Moss, & Daley, 1991). If we combine psychoactive substance abuse with alcohol use for people with schizophrenia, throughout life, this rate increases to 50% (Batel, 2000). One study carried out in America in the Epidemiological Catchment Area (ECA) estimated the current and lifetime occurrence of substance abuse among people with schizophrenia to be as high as 60% (Regier, Farmer, Rae, Myers, & et al., 1993; Swofford et al., 1996). A 12 month study by Swofford (1996) found substance use in schizophrenia patients to be a better predictor of relapse and hospitalization compared to fluphenazine decanoate dose. Other studies have also found AOD problems are related to schizophrenia relapse (Ayuso-Gutierrez & del Rio Vega, 1997).

11

Not surprisingly, ED is a core feature in schizophrenia (Gur et al., 2006). Research on the affect or emotion, experienced by individuals with schizophrenia, focus on either overt expression or physiological responses. Thus, research has mainly concentrated on emotional response as opposed to how one experiences difficulties in ER. A number of studies have reported that although individuals with the diagnosis of schizophrenia may experience similar levels of emotion when compared with controls, they have decreased emotion-expressive behaviour; in other words, they tend to outwardly display less emotion (Berenbaum & Oltmanns, 1992; Kring & Neale, 1996). Some studies have reported that people with schizophrenia experience more negative emotion in social interactions when compared to non-patient controls (Krause, Steimer-Krause, & Hufnagel, 1992).

There is also evidence that schizophrenia patients report experiencing the same amount of positive and negative emotion while viewing emotion-eliciting films or pictures compared to individuals without a diagnosis of schizophrenia (Berenbaum et al., 1992; Kring et al., 1996; Kring & Earnst, 1999). This indicates that people with a diagnosis of schizophrenia, while presenting with a 'flat affect' (a diminished outward expression of emotions (Andreasen & Olsen, 1982), are aware of, and can report on what internal emotions they are experiencing. However, few measures that specifically address difficulties in ER are utilised when researching this population (Berenbaum et al., 1992; Kring et al., 1996; Kring et al., 1999).

1.2 Why do we need a measure?

The majority of studies on ER tend to focus on coping strategies that are shown to be effective in reducing or regulating emotions. When emotions become a problem for the individual, clinicians tend to address the behaviour, either with pharmacology or counselling. While both address or pacify the emotions to some degree, a tool or measure of the difficulties in ER will enable clinicians to identify the degree of difficulty a client may experience with ER. Once identified, the clinician can address emotions specifically, as opposed to using interventions that focus largely on either changing or controlling cognitions or behaviours in the hope that emotions will regulate as a result. Also, having a scale that can detect the efficacy of a treatment intervention will provide the clinician with information regarding the client's progress during treatment. Furthermore, it will provide information regarding the outcome of treatment.

1.2.1 What does a measure indicate?

Evidence-based medicine has promoted the use of diagnostic screening scales to help guide clinical decision-making. These screening tools help indicate when a primary care patient should be classified as having a likely mental disorder and the degree to which they experience that disorder. Screening tools provide a link between clinical and community populations as they allow cross-disorder comparisons between physical and psychological health and assist in the provision of a comprehensive health service, planning information for decision makers, and importantly, treatment decisions that will enable clinicians to determine appropriate interventions. It is also part of evidence based practice to have a direct measure of the problem at the commencement of the therapy intervention and again at completion. This enables the clinician to determine the efficacy of therapy by measuring change, or to determine whether the correct therapy was implemented.

1.2.2 Multiple measures of constructs related to ER

Until recently there has been no single measure of ED for adults. This has meant that clinicians have had to use multiple measures that contain closely-related constructs in an attempt to capture the construct of ER (Mennin, Heimberg, Turk, & Fresco, 2005). The most frequently used measures of ER tend to refer to particular strategies that people may use to modulate emotional arousal while omitting other aspects of ER such as awareness, understanding and acceptance of emotions (Catanzaro & Mearns, 1990). One example is seen in a study by Garnefski and Kraaij (2006) who discussed cognitive ER strategies such as rumination, catastrophising, selfblame, other-blame, acceptance, positive reappraisal, putting into perspective, positive refocusing, and planning, to capture ER as opposed to difficulties in experiencing emotions. This study measured what someone thinks after their experience of threatening or stressful events using the Cognitive Emotion Regulation Questionnaire (CERQ) (Garnefski et al., 2006); a measure of depression symptoms (Symptom Checklist, SCL-90) (Derogatis, Rickels, & Rock, 1976) and a Likert scale (ranging from 1-5), to measure the use of cognitive strategies. The Trait Meta-Mood Scale (TMMS) (Salovey, Mayer, Goldman, Turvey, & Palfai, 1995) is another measure that is sometimes used to capture the construct of ER. Although this scale does tend to index emotional awareness, it fails to assess the ability to engage in desired behaviours when experiencing negative emotions.

Research that examined difficulties in regulating emotions among college students with and without generalised anxiety disorder used eight standardised questionnaires, none of which focussed specifically on difficulties in regulating emotions (Mennin et al., 2005). Rather, there were a cohort of tools designed to assess anxiety and to capture difficulties in regulating emotion. Some of these included: the Generalized Anxiety Disorder Questionnaire-IV (GAD-Q-IV) (Newman et al., 2002), the Affective Control Scale (ACS) (Williams, Chambless, & Ahrens, 1997) and, to assess the strength of emotional response tendencies, the Berkeley Expressivity Questionnaire (BEQ) was used (Gross & John, 1997). To assess worry the Penn State Worry Questionnaire (PSWQ) (Meyer, Miller, Metzger, & Borkovec, 1990) was administered, and the State-Trait Anxiety Inventory (STAI) (Spielberger, Gorsuch, Lushene, Vagg, & Jacobs, 1983) was used to assess chronic levels of anxiety. Symptoms of depression were assessed using the Beck Depression Inventory-II (BDI-II) (Beck, Steer, & Brown, 2000) and emotional intelligence was assessed with the Trait Meta-Mood Scale (TMMS) (Salovey et al., 1995) a scale that is related to mood and mood management.

Research by Leer-Felder, Zvolensky, Feldner and Lejuez (2004) used a measure that focussed on behaviour inhibition and emotional reactivity as opposed to difficulties in ER. This study discussed how generalised anxiety disorder (GAD) was associated with increased rumination (a form of emotional reactivity) and a type of ER. Measures in this study focussed on dispositional sensitivities in behavioural inhibition, negative effect, a response to situation scale that focussed on rumination, and a pictorial assessment of the affective meaning of the emotional experience. The study reported on dysfunctional style of rumination and regulating emotional distress as opposed to difficulties in ER.

Gross (1998b) proposed two processes of ER; antecedent-focused ER and response-ER. This study used a Likert-type scale to measure behaviour using a modified version of the Emotional Behavior Coding System (Gross & Levenson, 1997), and numerous measures of physiology. This study measured reactions to an event as opposed to an individual's general difficulty in ER.

The Self-Assessment Manikin (SAM) is a nonverbal graphic representation of emotional response. A number of manikin figures are presented displaying various emotions including happy (smiling figure) to unhappy (frowning figure). The client is then asked to rate a series of pictures that varied in both affective valence and intensity on a 9-point continuum (Lang, Greenwald, Bradley, & Hamm, 1993). This test is designed to measure emotional responses to a displayed situation.

Another measure of emotional responding is the Autonomic Nervous System (ANS) (Turpin, 1990). The ANS is popular as it does not rely upon the participant's subjective reporting, relying instead on physiological feedback such as skin conductance, cardiovascular activity and respiration. Similar to the SAM (Lang et al., 1993), it measures emotional response in relation to a current event or situation.

1.2.3 Problems with existing scales

Numerous scales are used to measure ER, however few measure difficulties in ER. Researchers tend to define and measure one or many single aspects of cognitive, expressive, behavioural and physiological processes. Aspects of ER that are generally measured can include one or more of the following - negative affect, that is, how an individual differs in their experience and expression of negative affect (Thomsen, Mehlsen, Viidik, Sommerlund, & Zachariae, 2005); or people's ability to rate the effectiveness of different strategies used for managing an emotionally charged situation; or the outward display of emotion (Kring, Smith, & Neale, 1994). Other aspects can include measuring the experience of emotion (Kring, Barrett, & Gard, 2003) or the ability to reflect on and monitor emotional responding (Salovey et al., 1995; Josephson, Singer, & Salovey, 1996). One study by Gross (2002) measured emotional expressivity, examining the behavioural (e.g. facial, vocal, postural) changes associated with the experience of emotion, such as smiling, laughing, frowning, storming out of a room, or crying. Another study by the same author examined emotional suppression (Gross & Levenson, 1993). Another study focussed on the degree of rumination one experienced (Nolen-Hoeksema, Morrow, & Fredrickson, 1993). Table 1-1 outlines

examples of instruments that have been used to report ER. As can been seen, although these measures report on certain aspects of ER, they fail to measure difficulties in ER.

Types of ER	Measure	Reports to measure ER*
Measure of emotional experience	Toronto Alexithymia Scale-20 (TAS- 20) (Bagby, Parker, & Taylor, 1994)	Assesses deficits in the cognitive processing of emotions and specifically the inability to accurately identify and label emotions
	Self-Assessment Manikin (SAM) (Lang et al., 1993)	Non-verbal graphic representation of emotional dimensions
	Autonomic Nervous System (ANS) (Turpin, 1990)	Biofeedback measure reflecting arousal responses
Measure of emotional expression/awareness	Levels of Emotion Awareness Scale (LEAS) (Lane, Quinlan, Schwartz, Walker, & et al., 1990)	Index of individual's emotional intelligence measures emotional awareness within a cognitive-developmental framework
	Trait Meta-Mood Scale (TMMS) (Salovey et al., 1995)	Measures emotional beliefs and attitudes people have towards their own emotional experiences
	The Berkeley Expressivity Questionnaire (BEQ) (Gross & John, 1995)	Relies on individuals' awareness of their level of expressiveness in order to accurately report on Likert-type scale
	Emotional Expressivity Scale (EES) (Kring et al., 1994)	Measures the extent to which an individual generally outwardly expresses positive and negative emotions
Measure of emotional strategies	Negative Mood Regulation Scale (Catanzaro et al., 1990)	A measure of generalized expectancies for negative mood regulation
Measure of emotional regulation	ER Questionnaire (ERQ) (Gross & John, 2003)	Measuring individual differences in expressive suppression and cognitive reappraisal

Table 1-1 Examples of instruments used to report emotional regulation/dysregulation

*ER: emotional regulation

A further problem arises because the majority of measures used to screen for ER were initially developed and validated on student populations as opposed to clinical populations. Student populations are a pragmatic and opportunistic resource for obtaining psychometric information. However, there is an assumption that student populations are representative of the broader population and cope with ER the same way as individuals with a clinical diagnosis. Evidence, however, suggests that constructs in the measures may not have the same meaning across groups, due to differences in demographics across populations, generations and socio-economic status, and can be contaminated by group-specific attributes that are unrelated to the construct of interest (Gregorich, 2006). Finally the psychometric properties of measures need to be reassessed in various groups including clinical and other cultural groups in an effort to ensure translational aspects of the measure. Translational aspects of a measure require it to be reliable and valid in a number of ways, for example from a student population to a clinical population; also from a healthy population to a clinical population. Furthermore, the measure needs to reproduce the psychometric outcomes across all groups, as reported in its initial development (Lambert & Hawkins, 2004).

This research will conceptualise the functionality of emotions using Gratz and Roemer's definition mentioned above (2004). The DERS was originally designed for individuals diagnosed with borderline personality disorder and showed sound reliability and validity (details of this will be discussed later in this Chapter). One specific behaviour, that of deliberate self-harm (the deliberate destruction or alteration of body tissue without conscious suicidal intent) (Gratz, 2001), is thought to occur among 70-75% of individuals with borderline personality disorder (Linehan, 1993; Gunderson, 2001) and is thought to be associated with the difficulties individuals experience in regulating emotion. This leaves approximately 30% of individuals with borderline personality disorder not engaging in self-harm and others who have the diagnosis yet can have episodes, or periods of time, where no self-harm occurs. While there can be no doubt that individuals diagnosed with borderline personality disorder experience difficulties in ER, there are numerous psychological disorders that involve disturbances in emotion and this is being recognised and measured by researchers.

1.3 Studies that report on the DERS

A number of studies have administered the DERS to participants experiencing a variety of disorders including, uncued panic attacks (Tull, 2006), generalized anxiety disorder, post-traumatic stress (PTS) (Tull, Barrett, McMillan, & Roemer, 2007), alcohol dependence (Fox, Hong, & Sinha, 2008), cocaine dependence (Fox, Axelrod, Paliwal, Sleeper, & Sinha, 2007), deliberate self-harm and substance use disorder (Martin, White, Flanagan, Yensel, & Bloomberg, 2011), pathological gambling (Williams, Grisham, Erskine, & Cassedy, 2012) and depression (Ehring, Fischer, Schnulle, Bosterling, & Tuschen-Caffier, 2008), and reported sound validity and reliability on the measure.

One study by Gratz et al (2006) focused on measuring change in symptoms over a 3 month period. This study offered 36 individuals with borderline personality disorder a partial hospitalisation program with a transition into an intensive outpatient program. Over half (56%) of patients reported a reliable improvement in ED with 44% reaching normal levels of functioning on the DERS.

The DERS was included in a study that examined individuals with borderline personality disorder and their willingness to tolerate emotional distress and difficulty engaging in goal-directed behavior when distressed. In this study the DERS was included to assess the construct validity of the experimental procedure and was expected to be negatively correlated with the Paced Auditory Serial Addition Task Computerized (PASAT-C) (Gratz, Rosenthal, Tull, Lejuez, & Gunderson, 2006). Findings supported the construct validity of the experimental procedure with ED negatively correlated with willingness to experience distress as measured by the PASAT-C.

An ER group intervention for deliberate self-harm among women with borderline personality disorder was the focus of another study that utilised the DERS. This program consisted of a 14-week program drawing from Acceptance and Commitment Therapy (Hayes, Strosahl, & Wilson, 1999), Dialectical Behavioural Therapy (Linehan, 1993), aspects of Emotion-focused psychotherapy (Greenberg, 2002) and traditional behaviour therapy. There were two conditions. The first group, received the group program in addition to their current outpatient therapy (N=12), while the second group continued with their current outpatient therapy alone (N=10). The DERS was employed in this study along with numerous other measures to give a baseline measure of ED. The DERS indicated significant levels of change on all subscales post a 14-week program (Gratz & Gunderson, 2006).

The DERS has also been used with a population who experienced uncued panic attacks (defined as discrete episodes of intense arousal and fear) (A.P.A., 2000) This study focused on the functional use of emotions according to Gratz's definition of ER (2004) and reported using emotionally avoidant regulation strategies (Tull & Roemer, 2007). There has also been evidence of deficits in ER associated with chronic worry and generalised anxiety disorder (GAD) (Salters-Pedneault, Roemer, Tull, Rucker, & Mennin, 2006). This study recruited 325 student participants, with 87 meeting the criteria for GAD. The non-clinical sample reported chronic worry associated with difficulties in ER in general and specific ER deficits across a number of domains (as measured by the DERS). Only the DERS subscale 'emotional awareness' was not found to be associated with chronic worry. Even after controlling for negative affect the relationship between analogue GAD and difficulties with a range of ER competencies (except for emotional awareness) were high (Salters-Pedneault et al., 2006). These latter two studies indicate that the DERS may be effective in measuring difficulties in ER in disorders other than borderline personality disorder.

A study examining post traumatic syndrome (PTS) in a sample of individuals who reported at least one traumatic incident in the past year found significant positive associations between PTS symptom severity and overall DERS scores, and on all subscales except the 'emotional awareness' subscale. This study also explored individuals who exhibited levels of PTS symptoms consistent with a PTSD diagnosis and found they scored higher on the DERS. These results remained significant after controlling for negative affect (Tull, Barrett et al., 2007).

The DERS has also been used to examine whether difficulties in ER are associated with early cocaine abstinence (Fox et al., 2007). This study compared healthy volunteers with treatment-seeking, cocaine-dependent individuals. Participant characteristics showed cocaine-dependent individuals differed significantly in alcohol use, and a high number met lifetime criteria for DSM-IV anxiety diagnoses including PTSD, however groups did not differ in mood disorder diagnoses. Cocaine-dependent individuals showed higher total DERS scores at baseline but there was no difference when compared with healthy controls at discharge. From baseline to discharge cocaine abusers reported significant improvement on the DERS total score, and the subscales of 'goals', 'strategies', and 'clarity'. There were no differences between baseline and discharge DERS scores on the subscales 'awareness', 'acceptance' and 'impulse control'.

One study examined ED, with the DERS among substance use disordered patients (SUD) with and without deliberate self-harm. It was found that ED remained significantly high among SUD patients with deliberate self harm even after controlling for the influence of other relevant risk factors and gender. Not surprising, SUD patients with deliberate self-harm reported higher levels of ED than SUD patients without deliberate self-harm. Higher level of overall ED were reported on the DERS subscales of 'nonacceptance', 'strategies' and 'goals' for the SUD with deliberate self-harm patients (Martin et al., 2011).

A study examining ED has also been found to be associated with pathological gambling. A study by Williams et al (2012) found pathological gamblers scored significantly higher on all DERS subscales, when compared to the healthy control group with the exception of the subscale 'goals'. The clinical control group (meeting the criteria for Axis I, mood and anxiety) scored significantly higher on all subscales compared to the pathological gamblers and healthy controls with the exception of the 'awareness' subscale, where pathological gamblers scored higher than either the clinical or healthy group.

Finally, the DERS was used in a study looking at the difficulties in ER and impulse control in recently abstinent people with alcohol dependence compared with social drinkers (Fox et al., 2008). This study compared 50 recently abstinent treatmentseeking alcohol dependent patients with social drinkers. Alcohol dependent patients reported an overall problem with difficulty regulating their emotions compared with social drinkers. Borderline personality disorder patients from the alcohol dependent sample reported differences on the DERS subscales of 'awareness' and 'impulse control' during the first week of treatment. Following protracted abstinence alcoholdependent patients without concurrent borderline personality disorder significantly improved on the subscales of 'awareness' and 'clarity' of their emotional experience. However, difficulties on the subscale of 'impulse control' persisted. This study showed the DERS' ability to discriminate between alcohol dependent and social drinkers as well as sensitivity to detect changes in ER during protracted alcohol abstinence indicating the DERS may be an appropriate tool for assessing ER construct in this population.

1.3.1 Limitations of studies using the DERS

There are several limitations associated with the above studies. The first being the use of undergraduate students as a population sample (Salters-Pedneault et al., 2006; Tull, Barrett et al., 2007; Tull & Roemer, 2007; Ehring et al., 2008). For example, Tull (2007) assessed students for incidents of panic and compared them to students who did not report an incident of panic; and Tull (2007b) also screened students who experienced an incident of trauma. Another study screened students who were currently non-depressed, however half had reported at least one major depressive episode in the past (recovered depressed group) and these were matched with a student control sample (never depressed group) (Ehring et al., 2008). The issue with this study is that, from a clinical perspective, recovered depression cannot be considered clinical depression. Also students are more likely to be high functioning individuals who are able to sit exams, meet deadlines for assignments and attend lecturers and tutorials weekly, therefore this form of screening does not represent a clinical population. It also makes it difficult to generalise results and outcomes to a clinical population.

Another limitation relates to sample size. Examples of the sample sizes used in these studies are presented in Table 1.2. Studies that did use a clinical population tended to have small sample sizes (n= 62 or as few as n=12), again making it difficult to generalise outcomes to the broader population being examined.

Author	Sample size	Clinical (DSM-IV)
Gratz, Tull and Gunderson (2008)	Group n=22 TAU n =12 TAU waitlist n=10	BPD
Gratz, Rosenthal et al (2006)	N= 18	BPD
	N= 18	PD
Gratz, Lacroce et al (2006)	Baseline, n=36 1 month, n=29 3 months n=18	BPD
Fox (2007)	N=60	Cocaine-dependent
	N= 50	Healthy volunteers
Fox (2008)	N=50	Alcohol dependent
	N=62	Social drinkers
Williams (2012)	N=56	Pathological gamblers
	N=50	Mixed clinical group Axis I (mood or
	N=49	anxiety.
		Healthy community controls
Staples and Mohlman (2012)	N= 37	GAD
	N=37	Non-GAD

 Table 1-2 An example of studies using the DERS with number of participants drawn from a clinical population

Lastly, while the psychometric properties of the DERS showed good validity and reliability, it was originally developed using an American student population. As such, its reliability and validity for use in Australian samples and in clinical populations is currently unknown. There have also been a number of studies that have highlighted inconsistencies in the subscales of the DERS suggesting that further psychometric validity of the measure is needed (Tull & Roemer, 2007; Weinberg et al., 2009; Bardeen et al., 2012).

1.3.2 Existing studies on the validity of the DERS

A number of studies have reported the psychometric properties of the DERS, however only a few have re-examined and published a factor analysis for the instrument. One study by Weinberg and Klonsky (2009) examined the psychometric properties of the DERS in a non-clinical adolescent population (aged 13-17 years) recruited from high schools. Results from this study supported the original six component factor structure, with sound reliability and validity. However, they also reported that six of the 36 items loaded

< .40 with a number of items cross loading or loading onto another factor. Despite this, they decided to score the DERS in their sample using all items from the original scale.

Weinberg (2009) also reported concerns regarding the six factor solution, noting that the 'awareness' subscale failed to correlate with five of the six clinical variables that were theoretically associated with ED. There could be numerous reasons for this lack of correlation including the scales use on a younger population, or that the language used in the items was not age-appropriate for adolescents. What this study highlights is that the scale may lack generalisability and that further research is needed on the psychometric properties of the scale to examine its appropriateness for use in different age, culture, or demographic populations.

Another study also reported problematic psychometric properties for the 'awareness' subscale (Tull & Roemer, 2007). Tull and Roemer conducted two studies and administered the DERS in a student non-clinical sample who were divided into two groups: experience of panic attack/s in the past year and no past history of panic attacks. Study one examined ER difficulties in the context of negative emotions or emotions in general. Study two examined ER difficulties experienced whilst watching a film clip and reporting reactions. Both studies reported that no difference was found between groups on the 'lack of emotional awareness' subscale and highlighted that the

22

items on the DERS may not adequately measure the specification of the model outlined by Gratz and Roemer (2004).

More recently, a study examined the latent structure of the DERS using a university student population. This study reported that once the 'awareness' subscale was removed, the revised five factor model was more optimal (Bardeen et al., 2012).

While comparison across studies, such as those described above, facilitate the use of the existing DERS measure, assessment of the psychometric properties of a measure across cultures and within clinical groups is also an essential first step (Van de Vijver & Leung, 1997). The importance of ensuring that a measure functions the same in other cultural settings allows researchers and clinicians to assess the similarities and differences in health impacts and performance across cultures and to identify potential translation differences. Furthermore, administering measures for the purpose of assessing outcomes in clinical practice makes it essential that the measure has acceptable reliability, validity, and can indicate sensitivity to change. This allows the clinician to accurately and reliably make decisions regarding treatment direction, and the efficacy of treatment (Hatfield & Ogles, 2004; Lambert et al., 2004). For these reasons further research on the stability of the DERS in a clinical population is essential.

1.4 Reliability and validity

1.4.1 Reliability

Reliability refers to the ability of an instrument to consistently measure a specific attribute (DeVon et al., 2007). In a general sense, something is considered reliable when the same result is gained over different occasions. For instance, if the car fuel gauge is not reliable it would be difficult to know when the tank is nearing empty and requiring a refill, or if a blood pressure gauge is not reliable it would be difficult to diagnose high or low blood pressure. It is similar when discussing the reliability of a instrument, in other words, for a instrument to be reliable, a person is required to obtain a consistent score when re-examined with the same instrument on a number of occasions.

Test-retest reliability is the method used to indicate the stability of an instrument. The time interval between the two administrations should be long enough

that participants do not remember their original responses, but not too long that fluctuations occur (Anastasi, 1990). High test-retest correlations indicate a more reliable scale. Computing Cronbach's coefficient alpha is also recommended when measuring the reliability of an instrument (DeVon et al., 2007). Cronbach's alpha provides a measure of the internal consistency of a scale and describes the extent to which all the items in a questionnaire measure the same concept. In general, if there is a low alpha value, it is due to poor correlation between items, and the item should be revised or discarded. On the other hand, if there are a number of items with a high alpha value above 0.90, it suggests that items are asking the same thing and are redundant. A minimum level of 0.7 Cronbach's alpha value is generally sought for instruments that contain more than ten items (Anastasi, 1990; DeVon et al., 2007).

1.4.1.1 The reliability of DERS reported in original study

The reliability of the DERS was originally assessed by Gratz (2004) using Cronbach's alpha to determine internal consistency of the DERS items. Results from 357 students indicated high internal consistency (a=.93) and item total correlations ranged from r=.16 to r=.69. Test-retest was conducted on 21 students and the overall DERS scores had sound test-retest reliability (p_i =.88, p < .01). The major limitations of this study include the small sample size for the test-retest and lack of a clinical population (2004).

Reliability is a prerequisite for validity and refers to the ability of an instrument to measure a particular trait or skill consistently. However, instruments can be highly reliable and still not be valid for a particular purpose.

1.4.2 Validity

Validity of an instrument refers to the extent the instrument measures what it purports to measure. For example, an instrument that is designed to measure performance is valid if its scores are directly related to future performance. Validity of an instrument is essential for the instrument to be used in decision making diagnosis, and treatment outcomes. Internal validity is the degree to which treatment makes change in the specific setting. There are a number of threats to internal validity including history, instrument change, statistical regression toward the mean, maturation, experimental mortality, selection and selection interaction (Wortman, 1983). External validity is the degree to which treatment effects can be generalised across populations, settings, treatment variables and measurement instruments (Dimitrov & Rumrill, 2003). Often the most common loss of external validity comes from experiments using volunteers, or employing small samples that are obtained from a single geographic location. It is necessary for a measure to have both internal and external validity. There are many different forms of validity, including: content validity, face validity, criterion-related validity (or predictive validity), construct validity, factorial validity, concurrent validity, convergent validity and divergent (or discriminant validity) (Gandek & Ware Jr., 1998). This thesis will re-examine and report on the criterion-related, factorial, concurrent and convergent validity of the DERS in later chapters.

1.4.2.1 The validity of DERS reported in original study

The original study conducted by Gratz and Roemer (2004) examined construct validity on the preliminary data provided by student self-reports, and compared this with another measure of ER which emphasized emotional avoidance, and expressive control, namely the Generalized Expectancy for Negative Mood Regulation Scale (NMR), (Catanzaro et al., 1990). It was reported that the overall DERS score was statistically significant and correlated with the NMR in the expected direction. Each of the DERS subscales was significantly correlated (in the expected direction) with the NMR and the measure of experiential avoidance. However, only three of the subscales were significantly correlated with the measure of emotional expressivity. When controlling for NMR, some overlap between these two scales was reported. However, this suggests that the NMR does not capture all clinically relevant aspects of ER. Predictive validity was also reported in the above study, which used 'liberated selfharm' (for example, 'the deliberate, direct destruction or alteration of body tissue without conscious suicidal intent but resulting in injury severe enough for tissue damage'), and frequency of intimate partner abuse. According to Gratz (2004) the limitations of this study including the small sample size for the test-retest, the large number of females (73 percent), and the lack of a clinical sample this suggests that the psychometric properties of the DERS need further examination.

1.5 Aims of research

This program of research has three specific aims.

- 1. To investigate the psychometric properties of the DERS in an Australian community population to determine if it can measure ER among Australian respondents in the same way as the original study (Gratz et al., 2004).
- 2. To re-examine the test-retest data reported in the Gratz (2004) study using a larger and more diverse sample. In the current thesis clinical respondents have been sourced from two populations: the first with an diagnosed Axis I disorder of Schizophrenia; the second with co-existing depression and AOD problems.
- 3. To assess the DERS' sensitivity to change pre-post intervention. This is important since it reflects the instrument's ability to detect change as a result of an intervention (Vermeersch, Lambert, & Burlingame, 2000; Lambert et al., 2004).

CHAPTER 2 DATA SOURCE

2.1 Overview

This Chapter describes the characteristics of four samples: 1) community, or control group, 2) a group of people with schizophrenia, 3) a group of people with coexisting depression and substance abuse (CDSA): and 4) a group of first year university students. All four samples will be used in this thesis to examine the psychometric properties of the Difficulties in ER (DERS) (Gratz et al., 2004).

The aim of this Chapter is to describe the demographics, symptomatology and characteristics, of participants in each sample. In addition, the recruitment strategies will be detailed in order to avoid repetition in later Chapters.

The first sample was drawn from the community; the second and third samples had a diagnosed mental illness and represented the clinical groups; namely schizophrenia and CDSA; and the fourth sample consisted of first year university students who were recruited at a later stage. The community sample (CS) completed the DERS at two time-points (test-retest). Both clinical samples completed the DERS at the initial stage, with the CDSA group also completing the DERS at 3, 6 and 12 month follow-up, post a 10 week treatment intervention.

Mean scores and standard deviations were generated and t-tests and analysis of variance (ANOVA) were conducted with each sample to examine the possible effects of age, education, and gender on the outcome scores on the DERS. Each sample had a total mean score on the DERS that reflected the degree of difficulty they experienced in ER with the control sample scoring lowest, indicating they experienced less difficulty in ER, while the CDSA sample scored above all other samples, suggesting they experienced a higher level of ER difficulties

2.2 Introduction

Emotional dysregulation (ED) is a central feature in many mental health disorders and has been implicated in over half of the DSM-IV Axis I and II disorders (James J. Gross et al., 1995; Hayes, Wilson, Gifford, Follette, & Strosahl, 1996; A.P.A., 2000; Bushman et al., 2001; Jukupcak et al., 2002; Briere, 2006). ED has also been linked to difficulties in an individual's ability to function across important domains of their lives including work and interpersonal relationships (Gross & Levenson, 1997). Chapter 1 has outlined the potential benefits of having a valid measure of difficulties in ER. Briefly, such a measure would guide a clinician's decision making in determining the most appropriate intervention to use with the client, and would provide an outcome measure to assist with determining the efficacy of the intervention.

The DERS (Gratz et al., 2004) was recently developed to measure the degree of difficulty in ER experienced by an individual. It is increasingly being used as a measure in numerous studies of mental health conditions such as panic attacks, (Tull & Roemer, 2007); generalized anxiety disorder, (Salters-Pedneault et al., 2006), post-traumatic stress disorder (Tull, Jakupcak, Paulson, & Gratz, 2007), alcohol-dependence, (Fox et al., 2008) and in cocaine-dependent individuals with depression (Fox et al., 2007) (refer to Chapter one for more details of these studies). While the original study by Gratz and Roemer (2004) reported good reliability and validity for the DERS using a student sample there were two major limitations: i) the findings required replication with a larger sample; and ii) the findings required cross-validation in a clinical sample. This thesis seeks to address these limitations, and to expand on the original study.

2.2.1 Summary of DERS measure

The DERS is a 36-item questionnaire designed to measure the difficulties an individual has in regulating emotions, otherwise referred to as ED. The items were chosen to reflect difficulties within four dimensions: (i) awareness and understanding of emotions; (ii) acceptance of emotions; (iii) ability to engage in goal-directed behaviour, and refrain from impulsive behaviour when experiencing negative emotions; and (iv) access to ER strategies perceived as effective (Gratz et al., 2004). Participants are asked to indicate how often the items apply to themselves, using a 5-point Likert-scale, where 1 is *almost never* (0-10% of occasions), 2 is *sometimes* (11-33%), 3 is *about half the time* (36-65%), 4 is *most of the time* (66-90%), and 5 is *almost always* (91-100%) (Appendix A-1).

The study by Gratz and Roemer (2004) recruited 357 university students of whom 73% were female. Exploratory factor analysis was used to provide preliminary data on the factor structure of the DERS and identified six underlying dimensions of ER that accounted for 56% of the total variance of the measured variables. The six interpretable factors were labelled: 'nonacceptance', with six items reflecting a tendency not to accept reactions to one's distress; 'goals', with five items reflecting difficulties concentrating and accomplishing tasks when experiencing negative emotions; 'impulse', with six items reflecting difficulties remaining in control of one's behaviour when experiencing negative emotions; 'awareness', with six items reflecting an lack of awareness of one's emotional responses; 'strategies', with eight items reflecting a belief that there is little one can do to regulate emotions once upset; and 'clarity' with five items reflecting an individual's knowledge about the emotions they are experiencing. High scores reflect greater difficulties in ER. Subscales are derived by using the mean score of the items, and have been found to have high internal consistency with *a* .80 for each.

To determine the test-retest reliability of DERS scores 21 participants were recruited by Gratz and Roemer (2004). The overall DERS score was found to have good reliability over a period ranging from four to eight weeks (.88, p<.01). The test- retest reliability of the DERS subscales was adequate, reporting inter-class correlations of .69 for 'nonacceptance', .69 for 'goals', .57 for 'impulse', .68 for 'awareness', .89 for 'strategies' and .80 for 'clarity'. A gender difference was found on the awareness subscale, with men reporting lower emotional 'awareness' than women.

Even though the DERS shows promise and was developed with regard to clinically relevant ER concepts, it may not perform as well when used in clinical populations. Testing on a community sample, especially a university student sample means that the factor structure may not be replicated when used in a clinical sample. Floyd and Widaman (1995) suggest 'failure to replicate across samples with known differences, such as clinical and nonclinical samples is generally due to the restriction of range of the variables within one or both of the samples'. It is therefore necessary to re-examine the psychometric properties of the DERS to determine whether it behaves similarly in a clinical sample when compared to that of a community sample.

2.2.2 Clinical samples

Two clinical samples were selected for the purpose of this research. The first sample consisted of people with a diagnosis of schizophrenia and the second sample was diagnosed with co-existing depression and substance abuse (CDSA). As outlined
in Chapter 1, both samples struggle with difficulties in ER, experiencing high levels of ED.

The first sample was included as it is widely known that individuals with schizophrenia are less expressive (both facially and vocally) than people without schizophrenia (Berenbaum et al., 1992; Kring et al., 1996). Previous research in people with schizophrenia suggested they experience a similar level of internal difficulty in ER when compared to others (Gruber & Kring, 2008), and can recognise and describe their own internal emotional experience when asked (Kring, Kerr, Smith, & Neale, 1993; Kring et al., 1996) (see Chapter 1). Previous research has predominately examined the external experience of emotion in this group such as the way they perceive and process emotion (Williams, Loughland, Gordon, & Davidson, 1999; Loughland, Williams, & Gordon, 2002), but has not yet adequately examined the internal experience of emotion in this group or the ability of people with schizophrenia to emotionally regulate, despite the observed emotional difficulties. One reason for this has been the lack of a sufficient tool for doing so. It would therefore be of scientific interest to gather further information on the experience of ED in people with schizophrenia using a reliable and valid self-report measure. This would provide detail regarding their internal experience as opposed to relying on clinical interviews, where individuals with schizophrenia may be less outwardly expressive (Kring et al., 1996).

It has been reported that individuals who abuse substances also complain of depressive symptoms at some time in their life. Raimo & Schuckit (1998) state as many as 80% of alcohol abusive individuals can report depressive symptoms. The CDSA sample in this thesis has been included for a number of reasons. First, all participants have been diagnosed with both depression and a substance disorder. It is widely reported that distressed individuals tend to indulge immediate impulses to make themselves feel better (Tice, Bratslavsky, & Baumeister, 2001). Second, impulse control has been identified as an important component of the addictive process (Evenden, 1999). Impulse control represents one of the major behavioural components of ER and is a subscale on the DERS (Gratz et al., 2004). The DERS has been used in two studies that examined substance abusing individuals and was recently validated in cocaine dependent patients (Fox et al., 2007; Fox et al., 2008). The CDSA sample has been

included in this research to examine the validity of the measure in individuals with a co-existing diagnosis of depression and substance abuse.

2.2.3 Community or control sample (CS)

In order to arrive at the most parsimonious factor solution using the DERS, a broad 'community' sample (CS) was included and consists of both University students and the general public. These two groups were combined to produce a combined CS, from which to compare the performance of the clinical groups (schizophrenia and CDSA) on the DERS. Both clinical samples lived in the community at the time of collecting data and could easily be targeted as part of the CS. The Australian Bureau of Statistics conducted a household adult survey in 1997 to establish a database of the distribution of mental disorders within the population and reported that from over 10,000 households that were sampled, 17.7% had one or more common mental disorders including anxiety, depression, alcohol or substance abuse and neurasthenia (Andrews, Hall, Teeson, & Henderson, 1999, 2000; Andrews, Henderson, & Hall, 2001). Hence, combining the clinical samples with the CS will more accurately reflect a sample that is representative of the general population and enable this study to determine the most appropriate composition of the subscales.

Chapter 3 will explore the psychometric properties of the DERS using testretest. However, the aim of this Chapter is to outline the characteristics of each sample comprising of the study participants.

2.3 Aims for Chapter Two

- 1. To describe recruitment, demographic characteristics, and the assessments used for each sample.
- 2. Report the mean scores and standard deviations on the basic demographic variables such as age, gender, marital status, education levels and employment and on the symptomatology variables for each sample.
- 3. To examine and compare the overall DERS mean scores and standard deviation scores found in each sample using t-tests in order to determine any difference in scores between the community and clinical samples.

4. To report on the mean and standard deviation scores of the combined sample: CS, Schizophrenia, and CDSA.

The CDSA sample will be used in this thesis with two aims:

- 1. First, it will be combined with the CS, and schizophrenia sample to establish psychometric properties of the DERS in Chapter 3.
- 2. Second, it will be used in Chapter 5 to examine whether the DERS can indicate change pre and post a treatment intervention.

Therefore, the additional aims for Chapter 2, for the CDSA sample, are as follows:

- 3. To report on the levels of depression and substance use in the CDSA sample using mean scores and standard deviations.
- 4. To determine if there is a difference in gender on the levels of depression and substance use.
- 5. To report on the relationship between type and severity of substance use and depression using ANOVA.

2.4 Methods

2.4.1 Participants and procedure

2.4.1.1 Community/Control Sample (CS)

Volunteers for CS were recruited from a number of sources including adult community and education groups in Newcastle, NSW, Australia, and students from various schools at the University of Newcastle, Australia. The sample was recruited between 2003-04, with a second phase between 2007-08. In order to collect data for the CS, the researcher approached Directors, and Presidents of local community organisations (such as the Lions and Rotary clubs) and educational organisations (such as adult community education centres) requesting permission to present to their members the purpose of the data collection, and to invite participants to complete questionnaires. The candidate also approached the Heads of Schools at the University of Newcastle and sought permission to enter scheduled lectures to present the research and invite participants to complete the questionnaires. To be eligible for this study CS participants were required to satisfy the following criteria:

- 1. Ability to understand and read English.
- 2. Have no existing or pre-existing diagnosis of a mental illness.
- 3. Be over the age of 18 years.

An information sheet was written specifically for the group being approached (Appendix B-1). It provided details about the nature of the research, the procedure and various contact numbers should any participant want to discuss the questionnaire, or make a complaint. The only detail that changed between recruitment places was information regarding how the questionnaire was collected/returned. This variation was in response to requests from the Director or President pertaining to that group.

In the initial assessment, participants completed the DERS and sociodemographic information, which took approximately 20 minutes. Participants were able to use a unique identification code known only to them which enabled data to be matched. A suggestion was provided for the identification in the information sheet (for example, mother's maiden name and day and month of birth). The questionnaire therefore did not identify individuals and consent was assumed when the completed questionnaire was returned. In order to gather re-test data (time 2) from the same individuals the researcher returned to the same location a week later requesting that participants complete the questionnaire a second time, using the same identification code allowing data to be matched.

Confidentiality for volunteers outside the university was maintained by handing out the re-test to all the people in the room. People were asked to leave the questionnaire package behind if they no longer wanted to participate or did not participate the previous week. Those unused packages were collected at the end of the class or meeting for re-use with other subjects. All completed questionnaires from the university students were returned via the school or free internal mail system which is accessible to students through their individual schools. This research was not a product of a research grant, and students did not receive any credit points or monetary payment for completing the questionnaires. By the end of 2007 a smaller than expected number of males had been recruited, prompting the researcher to approach the Hunter Medical Research Institute (HMRI). The HMRI maintains a database of volunteers from all over the Hunter Region who are interested in participating in research projects conducted by HMRI affiliated researchers. HMRI receives donations from large businesses and industry in the region. Many volunteers on the registry are employees from these supporting businesses and industry, creating a large database of volunteers for the purpose of research. An application was made to HMRI requesting access to males between the ages of 21-45 years who did not have a clinical diagnosis (Appendix A-2). This application was approved in March 2008. Questionnaires for this group were mailed out to each volunteer with written information and a postage paid self-addressed return envelope provided.

2.4.1.2 Diagnosed Schizophrenia sample

In order to access participants with a diagnosis of schizophrenia the researcher approached the Australian Schizophrenia Research Band (ASRB), previously referred to as NISAD, and applied to access volunteers enrolled on their 'Registry' database. Registry participants are recruited from a range of clinical settings, including mental health hospitals and community outpatient centres, non-government organisations such as the Schizophrenia Fellowship, doctor's surgeries, and through multi-media advertisements, promotional brochures, and events. All volunteers have a confirmed diagnosis of schizophrenia, using the Diagnostic Interview for Psychosis (DIP) (Jablensky et al., 1999), which is administered to all volunteers by a trained ASRB clinical staff member. A standard battery of assessment instruments are administered during the same session and basic demographic information, medical history, drug and alcohol usage are collected. An example of the additional assessments included: the National Adult Reading Test (NART) (Nelson, 1982); Repeatable Battery for neuropsychological assessment (RBANS) (Gold, Queern, & Bachanan, 1999); the Positive and Negative Syndrome Scale (PANSS) (Kay, Fiszbein, & Opler, 1986); and general functioning as measured by the Global Assessment of Functioning (GAF) (A.P.A., 1994). For a more comprehensive overview of the ASRB assessment procedure refer to the article by Loughland, Carr, and Lewin (2001; Loughland et al., 2010).

34

To be eligible for this study participants were required to satisfy the following criteria:

- 1. Ability to understand and read English.
- 2. Be over the age of 18 years.
- 3. Have a diagnosis of schizophrenia.
- 4. Not be actively unwell at the time of completing the questionnaire.

On approval from ASRB to access volunteers, the Registry posted letters to all ASRB members who met the eligibility criteria and lived locally to invite them to participate in the study (Appendix A-3).Volunteers were living in the community at the time of approach and were, therefore, considered to be psychiatrically well and sufficiently competent to consent or refuse to participate in research. Current research provides evidence that psychotic symptoms or having a poor capacity to make decisions, does not reflect an inability to understand information related to research participation (Carpenter et al., 2000). Reports have also indicated that even when assessments occur across changes in symptoms and medication status, individuals with schizophrenia provide self-report data on emotional experience that yield high internal consistency and test-retest reliability (Blanchard, Mueser, & Bellack, 1998; Horan, Green, Kring, & Nuechterlein, 2006; Kring & Moran, 2008).

Those individuals who provided consent to participate in the present study were posted a package containing demographic questions plus the DERS, the Kessler-10 (Andrews & Slade, 2001), and the Social Problem-Solving Inventory-Revised (Thomas J. D'Zurilla et al., 1998), along with pre-paid self-addressed return envelopes. The researcher was not provided with access to the volunteers directly and all communication was passed between the ASRB Registry and the participant. The ASRB Registry de-identified all returned questionnaires prior to returning them to the researcher.

2.4.1.3 Co-existing depression and substance abuse (CDSA)

Participants in this thesis were sourced from a larger existing study referred to by the term SHADE (<u>Self-Help</u> for <u>Alcohol/other</u> drugs and <u>Depression</u>) which was funded in part by the Australian Education and Research Foundation (AERF). For a comprehensive report on the intervention, and therapist's adherence to protocols refer to studies by Kay-Lambkin et al (Kay-Lambkin et al., 2009; Kay-Lambkin, Baker, Lewin, & Carr, 2011; Kay-Lambkin, Baker, Kelly, & Lewin, 2011). Participants for SHADE were recruited from a range of settings in the Hunter and Greater Western Regions of New South Wales, Australia. Referrals to the project were received from drug and alcohol clinical services, mental health services, primary health care settings, and the general community (this latter group responded to local advertising).

To be eligible for the SHADE study, participants were required to satisfy the following criteria:

- Current depressive symptomatology of 17 or greater on the Beck Depression Inventory-II (BDI-II) (Beck, Steer, & Brown, 1996).
- Lifetime diagnosis of major depressive disorder, as confirmed by the Structured Clinical Interview for DSM-IV, Research Version (SCID-RV) (First, Spitzer, Gibbon, & Williams, 2001);
- 3. Current intake of at least one of the following: alcohol (consumption above recommended safe drinking levels as suggested by the Australian National Health and Medical Research Council (NHMRC, at the time of assessment in 2009); equating to four standard drinks per day for men or two standard drinks per day for women with fewer than two alcohol free days per week); cannabis (at least weekly use); or amphetamines (at least weekly use);
- 4. Absence of a brain injury, organic brain disease and/or significant cognitive impairment;
- 5. Ability to understand English; and
- 6. Be over the age of 18 years.

SHADE participants were assessed for eligibility and randomised into one of three interventions. Following randomisation, and completion of a 10-week treatment intervention, participants were reassessed by independent research assistants blind to intervention allocation. A battery of questionnaires were administered as part of the SHADE study and will be outlined below. Details regarding the interventions and other aspects of SHADE are described in Chapter 5.

The DERS was included in the SHADE study at four stages: baseline, 3- 6- and 12 months follow-up. The aim of using this sample was twofold: (i) to examine the

validity of the DERS in a clinical group that have a dual diagnosis; and (ii) data to examine whether the DERS shows change post treatment intervention, the findings of which are presented in Chapter 5. Participants from the SHADE study that are employed in this thesis are referred to as the CDSA sample.

2.4.1.4 University student sample

First year psychology students were offered two hours of course credit in exchange for volunteering to complete a series of questionnaires that were part of a larger study conducted by Dr Miles Bore, of the University of Newcastle, Australia, and titled 'Personality Structure and Psychological Health'. This sample was recruited later than the other samples in order to provide an additional data set to assess construct validity, reported in Chapter 5.

To be eligible for this study participants were required to satisfy the following criteria:

- 1. Ability to understand and read English.
- 2. Be over the age of 18 years.
- 3. Be enrolled in first year psychology studies.

The assessment tools used for the purpose of this study were the Difficulties in Emotion Regulation Scale (DERS) (Gratz et al., 2004), Kessler 10 (K10) (2002) and Goldberg's International Personality Item Pool (IPIP) (1992).

2.4.2 Assessment instruments

All four samples described above were administered the DERS and general demographic information was collected.

2.4.2.1 Demographic information

Basic demographic characteristics were collected over the following domains: age, gender, education, work status, and relationship status. Recognising that individuals in either a student or community population might also have a diagnosis of mental illness, an additional question asked participants to tick a box if they had ever been treated for a mental illness. This question enabled the researcher to exclude those with a diagnosed mental illness from the CS (Appendix B-1).

2.4.2.2 Difficulties in ER Scale (DERS) (Gratz et al., 2004)

The DERS is a 36 item self report measure and has been described in the introduction of this Chapter. The DERS scale demonstrated high internal consistency overall and good internal consistency for all of the six factors (reported above). The reliability and validity of the DERS will be examined in Chapter 3; construct validity in Chapter 4, and finally, an examination of whether the DERS can detect change after a treatment intervention in Chapter 5.

2.4.2.3 Social Problem Solving Inventory (SPSI-R) (Thomas J. D'Zurilla et al., 1998)

SPSI-R is a 52-item multidimensional self-report measure of social problemsolving ability that accesses problem-solving skills and problem orientation. Each item is rated on a 5-point scale ranging from 'not at all true of me' (0) 'to extremely true of me' (4) by the respondent. The SPSI-R consists of five unidimensional scales: positive problem orientation (PPO; 5 items), negative problem orientation (NPO; 10 items), rational problem solving (RPO; 20 items), impulsivity/carelessness styles (ICS; 10 items) and avoidance style (AS; 7 items).

The measure has been shown to have adequate reliability and validity among adults with test-retest reliability (3 week period) ranging from 0.72 (PPO) to 0.88 (NPO) in a sample of 138 college students (D'Zurilla, Chang, Nottingham, & Lino, 1998). In second test-retest study coefficients were reported to ranged from 0.68 (PPO) to .091 (NPO) in 221 nursing students (Thomas J. D'Zurilla et al., 1998). The SPSI-R has shown to be useful in predicting depression (Nezu & Ronan, 1985; Priester & Clum, 1993), hopelessness (Schotte & Clum, 1982) and suicidal ideation (Dixon & Heppner, 1987). It takes 15-20 minutes to complete. It will be used in Chapter 4 to examine construct validity of the DERS.

2.4.2.4 Kessler 10 (K10) (Kessler et al., 2002).

The K10 has been shown to be a sensitive screen for CIDI/DSM-IV disorders (Ayuso-Gutierrez et al., 1997; A.P.A., 2000) in surveys carried out in the US (Kessler et al., 2002). It is a 10-item self-report questionnaire designed to measure non-specific psychosocial distress based on questions about levels of nervousness, agitation, psychological fatigue and depression. Questions ask respondents how frequently they

experienced symptoms of psychological distress (eg, feeling so sad that nothing can cheer you up) during the past 30 days, past 12 months and so forth.

The K10 has been used widely, and was included in the World Health Organization World Mental Health Survey as a clinical outcome measure (W.H.O., 1997). The factorial composition of the K10 was examined in a prospective community survey and cross-validated in a separate large community survey (Andrews, Sanderson, & Beard, 1998). Each item is scored from 1-5 from "none of the time" to "all of the time", with higher scores indicating greater distress. There are four additional items that do not contribute to the total score and serve to assess the degree of disability associated with the individuals' psychological distress.

The four factors of the K10, labelled 'Nervous', 'Negative Affect', 'Fatigue', and 'Agitation', were consistent with the original scales from which the items were taken. (Brooks, Beard, & Steel, 2006). Andrews & Slade (2001) have developed normative data using the Australian National Survey of Mental Health and Well-Being (NSMHWB) data conducted in 1997. The K10 performed well compared to other the mental health instruments, such as the General Health Questionnaire (GHQ) (Goldberg & Williams, 1988), the quality of life instrument (SF-12) (Kessler et al., 2002); and current diagnosis of anxiety and affective disorders according to the Composite International Diagnostic Interview (CIDI) (Furukawa, Kessler, Slade, & Andrews, 2003). In a study by Furukawa et al (2003) the K10 out performed both the K6 and GHQ in screening for CIDI/DSM-IV, mood and anxiety disorders (W.H.O., 1997; A.P.A., 2000). It is included in this study to examine the construct validity of the DERS and will be reported in Chapter 4.

2.4.3 Additional assessment Instruments employed from the SHADE study

Participants in the CDSA clinical sample were required to complete a battery of questionnaires as part of the SHADE project in order to establish that participants met the eligibility criteria. Once participants were accepted into the SHADE project the DERS was included with the SHADE battery of assessments, for the purpose of this study.

Questionnaires for the SHADE project included in this thesis are:

2.4.3.1 Alcohol Use Disorders Identification Test (AUDIT) (Babor, Higgins-Biddle, Saunders, & Monteiro, 2001)

The AUDIT contains 10 multiple choice questions on the quantity and frequency of alcohol consumption, drinking behaviour and alcohol-related problems or reactions. It focuses on identifying the preliminary signs of hazardous drinking and dependence and is used to detect alcohol problems experienced in the previous 12 months. Questions 1-3 deal with alcohol consumption, 4-6 relate to alcohol dependence and 7-10 consider alcohol related problems. A score of 8 or more in men (7 or more in women) indicates a strong likelihood of hazardous alcohol consumption. A score of 13 or more for women and 15 or more for men is suggestive of alcohol dependence (Saunders, Aasland, Babor, de le Fuente, & Grant, 1993; Babor et al., 2001)

2.4.3.2 The Drug Use Scale of the Opiate Treatment Index (OTI) (Darke, Ward, Hall, Heather, & Wodak, 1991)

The Drug Use Scale of the OTI reports on the quantity and frequency of substance use across 11 drug types including: alcohol, cannabis, heroin, other opiates, amphetamines, cocaine, hallucinogens, barbiturates, tranquilisers, inhalants and tobacco. Each drug type is individually assessed and clients report on the last three occasions of use in the month prior to assessment, estimating the amount of drug consumed on each of these occasions. The scales on the OTI can be used as a whole or individually without compromising the validity or reliability of the scale. Each score represents the average occasions of use per day, for the previous month.

2.4.3.3 Structured Clinical Interview for DSM-IV (SCID) (First et al., 2001)

The SCID is a semi-structured interview and provides a diagnostic clinicianrated measure of alcohol abuse and dependence based on the criteria set in DSM-IV (A.P.A., 2000). It relies in part on clinician judgement to derive a diagnosis. Therefore, reliability of the scale is related to the context in which it is being used (First et al., 2001). In order to reduce clinician error, an intensive training program took place, using video based case studies, where clinicians would rate cases and a discussion of individuals' ratings, including rationale for discrepant ratings. This training maximized inter-rater reliability for the SHADE study.

2.4.3.4 Beck Depression Inventory-II (BDI-II) (Beck et al., 1996)

The BDI-II is commonly used to screen for depressive symptoms among people with drug and alcohol use problems (Dawe, Loxton, Hides, Kavanagh, & Mattick, 2002). It is a 21-item self-report questionnaire used to screen for depressive symptoms over the previous two-week period, and takes around 10 minutes to complete. The BDI-II has good internal consistency among psychiatric outpatients (*a*=0.93) and with a community sample (α =0.93). Test-retest reliability scores are high in psychiatric and non psychiatric samples (Beck et al., 1996). Scores range from 0 – 63, with the following cut-off points indicative of varying levels of severity of depression: 0-13 minimal depression; 14-19 mild depression; 20-28 moderate depression; 29 and over severe depression. High scores on the BDI-II indicate the severity of depressed mood. Beck and colleagues (Beck et al., 1996) recommend that respondents scoring 17 points or higher be referred for further assessment for major depressive disorder.

2.4.4 Additional measures from the university student study

Participants in the student sample were required to complete a battery of questionnaires as part of a larger project. For the purpose of this study students completed the DERS, the K10 and the IPIP.

2.4.4.1 Goldberg's International Personality Item Pool (IPIP)(Goldberg, 1992)

The IPIP Big-Five factor markers consist of a 50-item inventory which can be freely downloaded from the internet. The items have been translated into more than ten languages. It is comprised of five personality factors and their characteristic traits are described as 'neuroticism' which refers to worried, insecure, nervous, and highly strung; 'extroversion' described as sociable, talkative, fun-loving and affectionate; 'openness' described as original, independent, creative, daring; 'agreeableness', indicated by being good-natured, soft-hearted, trusting and courteous; and 'conscientiousness' indicated by careful, reliable, hardworking and organised stability (McCrae & Costa, 2004). Each item is scored from 1-5 with responses 'very inaccurate' assigned a value of 1, 'moderately inaccurate', 'neither inaccurate or accurate', 'moderately accurate', and 'very accurate', all values are summed to obtain a total scale score. Goldberg conducted a number of studies to examine the psychometric properties and found the Big-Five factors emerged distinctly across four analyses (Goldberg, 1992). Another recent study examined the psychometric properties of the IPIP using three different adult samples (n=906) and found the 5-factor solution in each sample to have good internal consistency and to relate strongly to major dimensions of personality assessed by the NEO Five Factor Inventory (NEO-FFI) and the Eysenck Personality Questionnaire-Revised Short Form (EPQ-R) (Gow, Whiteman, Pattie, & Deary, 2005). The IPIP has been included in this study to examine the construct validity of the DERS and will be reported in Chapter 4. Table 2-1 shows a summary outline of the questionnaires administered to each sample.

Instruments	Control/ community sample	Schizophrenia sample	CDSA sample	Student Sample
Demographic Information				\checkmark
DERS	\checkmark	\checkmark	\checkmark	\checkmark
SPSI-R	\checkmark	\checkmark	-	-
Kessler 10	\checkmark	\checkmark	-	\checkmark
BDI-II	-	-	\checkmark	-
OTI	-	-	\checkmark	-
AUDIT	-	-	\checkmark	-
SCID	-	-	\checkmark	-
IPIP	-	-	-	\checkmark

Table 2-1 Questionnaires administered to each sample, summary

2.4.5 Ethics

These procedures were carried out in accordance with the NHMRC's Statement of Ethical Conduct of Research among Human Participants. Ethics approval to recruit the community and schizophrenia participants was gained from the Human Research Ethics Committee at the University of Newcastle (HREC Approval No: H-777-0304). In order to access individuals with a confirmed diagnosis of schizophrenia a second application was made to the Schizophrenia Research Institute (SRI) formally known as the Neuroscience Institute of Schizophrenia and Allied Disorders (NISAD) in 2004. This was approved on 3rd May, 2004 (PO02/04). In 2008 a second application was submitted to HREC, as the three year period for collecting data had expired, and additional male recruits were required (HREC Approval No. H-2008-0204). In order to recruit males in the workforce an additional application was submitted to the Hunter Volunteer Register (Hunter Medical Research Institute) and was approved in March 2008. Ethics approval for the CDSA was gained from both the Human Research Ethics Committee at the University of Newcastle (HREC Approval No: H-750-0204) and the Hunter Area Research Ethics Committee (HAREC 03/12/10/3.17). Finally, a university sample was recruited in 2009 in order to provide additional data to assess construct validity, (HREC Ethics Approval No. H-2009-0089), (refer to table summary provided in Appendix C).

2.4.6 Statistical analysis

Data were analysed using the Statistical Package for Social Sciences (SPSS) for Windows, Version 16.0. Responses to the battery of questionnaires were collated and five data samples were formed:

- 1. Community or baseline, time 1 (CS, n= 194), community, time 2 (n= 168).
- 2. Two clinical samples; schizophrenia (n=125) and
- 3. Coexisting depression and substance abuse (n=231);
- A combined data set comprising of the CS, time 1, and both clinical sets (n=524)
- 5. A student only sample (n=237).

Each variable was checked for errors or scores out-of-range using frequency graphs before total scores were calculated. Frequency statistics were used to examine categorical variables and descriptive statistics were used to describe continuous variables.

For calculating scores on the DERS, first negatively worded item scores were reversed, then mean and standard deviations were calculated. T-tests were used to examine if a difference on DERS scores occurred as a result of gender. T-test and ANOVA were also used to compare the total DERS scores in each sample, and to compare age, education and work. An overall DERS mean score was not reported in the Gratz & Roemer study (2004) however a separate DERS mean for women and men was reported. Therefore to calculate an overall mean score to reflect all the samples, a mean was calculated using the mean from both the women and men [($\underline{m} \times 260$) + ($\underline{m} \times 97$) ÷ 357] and entered into a one-sample t-test with the DERS mean from the CS (time 1).

2.5 Results

2.5.1 Characteristics of the community sample

Exploratory data analyses were conducted on the CS, (time 1) and basic demographic variables such as age, gender, marital status, education levels and employment are reported. As seen in Table 2-2, 198 people were recruited to the current study, with 37% between the age of 18-30 years. Of these, 62.6% (n=124) were female and over half were married or living in a defacto relationship. The majority were either students attending university (31.8%) or had completed a university degree (44.9%), 46% in the age range of 18 - 25 years. A high majority were in the workforce n=158, of which 37.6% were part-time employed (see Table 2-2).

		Participants n=198	
		n=198	%
Gender	Males	74	37.4
	Females	124	62.6
Relationship	Single, never married	68	34.9
status	Married/Defacto	107	54.9
	Separated/Divorced/Widowed	20	10.3
Work Situation	Full-time	85	43.8
	Part-time	73	37.6
	Volunteer	7	3.6
	Unemployed	29	14.9
Education	Year 8 - 11	15	7.6
	HSC/TAFE/Trade	31	15.7
	Attending University	63	31.8
	Awarded University Degrees	89	44.9
Age range	18-30	73	37
	31-40	56	29
	>40	67	34
Levels of Difficulties in ER, n=196		Mean	SD
DERS Total		66.53	18.21
Nonacceptance		10.71	4.21
Goals		12.53	4.48
Impulse		8.90	3.32
Strategies		12.89	5.12
Aware		12.63	3.74
Clarity		8.85	2.72

Table 2-2 Demographic characteristics of the community sample (CS).

*Where subtotals <198, missing data due to self administration.

The total mean score for the DERS was 66.53 (SD=18.21), with means ranging for the six subscales between 8.85 – 12.89. Further exploration of the DERS score and subscale scores will be reported in Chapter 3.

2.5.2 Characteristics of clinical samples

2.5.2.1 Schizophrenia sample

One hundred and twenty-nine participants with a diagnosis of schizophrenia, were recruited to the current study, with 80% over the age of 31 years. Table 2-3 shows there was a similar number of males and females with 51.2% (n=66) male. As seen in Table 2.3 the sample was mostly single, 62.5% (n=80), with 37.5% either married, living in a defacto relationship or divorced. Nearly half of the sample were unemployed (43.8%). However, 56.2% were either working full-time, part-time or in a volunteer

capacity. Fifty-one percent had completed either the HSC, or TAFE, or had a trade qualification.

Total mean for the DERS was 87.40 (SD=23.94), with mean range for six subscales between 10.82 - 19.30. A difference of 21 in the DERS mean scores between this clinical sample and the control sample may indicate the DERS is able to detect a difference between the two samples. An independent t-test was conducted between this clinical sample and the control sample on the DERS mean scores and was found to be statistically significant (t (321) =-8.27), p<0.00, with participants in the schizophrenia sample experiencing higher difficulties in ER.

		Participants		
		n=129	%	
Gender	Males	66	51.2	
	Females	63	48.8	
Relationship	Single, never married	80	62.5	
status	Married/Defacto	28	21.9	
	Separated/Divorced/Widowed	20	15.6	
Work Situation	Full-time	14	10.9	
	Part-time	37	28.9	
	Volunteer	21	16.4	
	Unemployed	56	43.8	
Education	Year 8 - 11	28	22.2	
	HSC/TAFE/Trade	65	51.6	
	Attending University	10	7.9	
	Awarded University Degrees	23	18.3	
Age range	18-30	20	16.0	
	31-40	52	41.6	
	>40	53	42.4	
Levels of Difficulties in ER, n=125		Mean	SD	
DERS Total		87.04	23.94	
Nonacceptance		14.11	6.10	
Goals		15.42	4.23	
Impulse		12.77	4.83	
Strategies		19.03	7.24	
Aware		14.86	4.42	
Clarity		10.82	4.21	

	Table 2-3 Demographic	characteristics	of the schizo	phrenia sample
--	------------------------------	-----------------	---------------	----------------

*Where subtotals <129, missing data due to self administration

2.5.2.2 Co-existing depression and substance abuse sample

Two hundred and thirty-one participants with a diagnoses of CDSA were recruited to the current study, 50% over the age of 40. As seen in Table 2-4, just over half were male (n=129, 55.8%) and unemployed (n=117, 51.1%). Most of the sample (79%) were over 30 years of age. According to the guidelines for interpretation of scores on the BDI-II, a score of 17 or above indicates nominal threshold above which participants are deemed to have moderate levels of depression (Beck et al., 2000).

		Partic	ipants
		n=231	%
Gender	Males	129	55.8
	Females	102	44.2
Relationship	Single, never married	86	37.2
status	Married/Defacto	74	32.0
	Separated/Divorced/Widowed	71	30.7
Work Situation	Full-time	45	19.7
	Part-time	37	16.2
	Volunteer	4	1.7
	Unemployed	117	51.1
	Not in workforce	26	11.4
Education	Year 8 - 11	63	27.5
	HSCTAFE/Trade	148	64.6
	Attending University	0	0
	Awarded University Degree	18	7.9
Age range	18-30	46	20.0
	31-40	66	28.7
	>40	118	51.3
Levels of Difficulties in ER, n=229.		Mean	SD
DERS Total		108.95	21.47
Nonacceptance		17.25	5.67
Goals		17.53	4.23
Impulse		16.58	5.73
Strategies		24.67	6.67
Aware		18.29	5.10
Clarity		14.68	4.21

Table 2-4 Demographic characteristics of CDSA sample

Where subtotals <231, missing data due to self administration.

All participants in CDSA met criteria for moderate to severe depressive symptoms at the time of assessment with an overall mean =32.54, SD = 9.88. The mean was higher for females compared to males (males M =30.77, SD =8.27; females M=34.70,

SD=11.00). There was a significant difference between gender and depression levels F(1, 229) = 7.43, p< 0.00.

The guidelines used in the SHADE trial recommended that no more than four standard drinks per day for men or two standard drinks for women with at least two alcohol free days per week. The threshold for cannabis for inclusion was set at least weekly use, and for amphetamines at least weekly use (Kay-Lambkin et al., 2009). Table 2-5 also displays the rates at which participants in the study met the criteria for current substance use status as measured by the OTI. According to the scores on the OTI, participants who meet 'above threshold' for substance abuse were consuming more than eight standard drinks per day (M=8.23, SD=8.83), smoking cannabis on average five times per day (M=4.71, SD=11.03), and were using amphetamines once per week (M=0.16, SD=0.83).

A total DERS mean difference of 21 between the schizophrenia sample and CDSA in this study sample may indicate the DERS is able to detect a difference between the two clinical samples. A two sample t-test between the two clinical samples, overall DERS mean scores were found to be statistically significant t (428)=-22.28, p < 0.00 with participants in CDSA sample experiencing higher difficulties in ER.

	Mean Depression			
Above threshold	n=231	Scores	SD	
Alcohol Only (ALC)	110	31.25	9.03	
Cannabis Only (CAN)	44	34.11	12.50	
Amphetamines Only (AMP)	3	41.33	10.30	
ALC and CAN	41	34.85	11.90	
ALC and AMP	1	39.00	-	
CAN and AMP	7	30.29	7.30	
3 substances (ALC, CAN, AMP)	7	35.29	10.40	
Total	211	32.80	10.50	

Table 2-5	Depression	score with	substance	use
-----------	------------	------------	-----------	-----

Above threshold refers to depression scores greater than 20, according to BDI-II.

Above threshold indicates use of that substance is above recommended consumption rates.

Where subtotals <231, missing data due to self administration.

All individuals in this study used substances and all were above the threshold for depression (>17) on the BDI-II scale. Table 2-5 shows all individuals were above threshold for at least one type of substance use. There was also no significant difference between depression levels and the type of substance used, F(6, 205)=1.310, p< 0.25. In other words, depression levels remained high regardless of the type of substance consumed.

	n=231	Threshold	BDI-II Mean	SD		
Alcohol	22	Abstinent	35.23	9.31	F(2, 227),=0.71, p<0.51	
	48	Below	32.31	12.42		
	160	Above	32.51	9.98		
Cannabis	120	Abstinent	31.48	9.31	F(2, 229),=2.30, p<0.10	
	12	Below	31.08	8.59		
	100	Above	34.41	11.71		
Amphetamine	204	Abstinent	32.57	10.61	F(2, 228),=0.51, p<0.60	
	9	Below	30.4	6.98		
	18	Above	34.56	10.41		
Cannabis Amphetamine	120 12 100 204 9 18	Abstinent Below Above Abstinent Below Above	32.51 31.48 31.08 34.41 32.57 30.4 34.56	9.98 9.31 8.59 11.71 10.61 6.98 10.41	F(2, 229),=2.30, p<0.10 F(2, 228),=0.51, p<0.60	_

Table 2-6 Depression and substance severity

Abstinent indicates no use of nominated substance

*Using below indicates use of that substance is below recommended consumption rates or threshold for inclusion into the study

*Using above indicates use of that substance is above recommended consumption rates or threshold for inclusion into the study

*Where subtotals <231, missing data due to self administration

Table 2-6 displays depression status according to severity of substance use. Again depression levels remained high regardless of whether participants were abstinent, below or above the use threshold for a particular substance. A one-way ANOVA indicated scores on depression remain high, and this was not indicative of severity of substance use.





Figure 2:1 shows the lifetime use of substances by individuals was high, with nearly all participants having used alcohol (99.6%), and over 89% of the sample having used tobacco at some stage in the past. Over 59% of the sample had previously used amphetamines, 49% had used hallucinogens, and 32% had used cocaine. Also, the majority of the participants had used more than one substance.

2.5.2.3 Characteristics of the combined data sample: CS, Schizophrenia and CDSA

The combined sample included all data from the CS, schizophrenia and CDSA, and consisted of 558 participants, with 43% over the age of 40. There was a similar number of males and females with 48% (n=269) males. As seen in Table 2.7 the sample was mostly single (42%, n=234). Over half the participants (58%) were in paid employment with 80.8% of participants gaining an educational qualification above high school level. Total mean for the DERS was 88.00 (SD=23.94), with mean range for six subscales between 11.18 - 15.52.

		Participants		
		n=558	%	
Gender	Males	269	48.2	
	Females	289	51.8	
Relationship Status	Single	234	42.2	
	Never married/Defacto	209	37.7	
	Separated/Divorced/Widowed	111	20.0	
Work	Full-time	144	26.1	
	Part-time	147	26.7	
	Volunteer	32	5.8	
	Unemployed	202	36.7	
Education	Year 8 - 11	106	19.2	
	HSC/ TAFE/Trade	244	44.1	
	Attending University	73	13.2	
	Awarded University Degrees	130	23.5	
Age	18-30	139	25.2	
	31-40	174	31.6	
	>40	238	43.2	
Levels of Difficulties in ER		Mean	SD	
Total DERS		88.00	23.94	
Nonacceptance		14.20	6.03	
Goals		15.26	4.84	
Impulse		12.97	5.85	
Strategies		11.18	8.16	
Aware		15.52	5.21	
Clarity		11.71	4.55	

Table 2-7 Demographic characteristics for the combined community and clinical samples

*Where subtotals <558, missing data due to self administration

2.5.2.4 Characteristics of the university student sample

The student sample consisted of 264 participants, with an average age of 18 years (17-19 years = 48.1%). Most were female (168, 78.5%). As seen in Table 2-8 all were first year university students. Total mean on the DERS was 85.14 (SD=22.57), with the mean range for the six subscales between 11.29 -24.00.

		Participants		
		n=264	%	
Gender	Males	37	15.6%	
	Females	186	78.5%	
Education	Attending university	264	100%	
Age range	17-20	133	56.1%	
	21-53	104	43.9%	
Levels of Difficulties in ER		Mean	SD	
DERS Total		85.14	22.57	
Nonacceptance		13.88	5.92	
Goals		15.91	4.50	
Impulse		24.00	12.32	
Strategies		17.66	7.06	
Aware		14.24	4.48	
Clarity		11.29	3.83	

Table 2-8 Demographic characteristics for the university student sample,

*Where subtotals <264, missing data due to self administration

2.6 Discussion

The aim of this Chapter was to describe recruitment and characteristics of samples that will be used to examine the psychometric properties of the DERS in later Chapters. The CS in this study were drawn from a range of areas including university students, local registered clubs, and the general workforce, making this sample representative of the general community as opposed to the student only sample reported in the original study (Gratz et al., 2004). A difference in gender was found between CS in this sample having 62% female compared to Gratz and Roemer reporting 73% female, and in relationship status with 55% in this study being in a current relationship compared to 89% being single in the published study.

The Gratz and Roemer study (2004) did not report on work status or education except to say there was little difference demographically between participants who completed all measures, making it difficult to compare these variables with the present study. In the CS sample 44% were working full-time and 38% were working part-time. Individuals were also above average in education with 45% awarded a university degree, and 31% attending university. The characteristics reported in the CS vary from the university student sample used to develop the DERS measure as the total DERS mean score in the CS (M=66.4) was lower than the total DERS mean score in the Gratz and Roemer study (M=78.7). This indicates that participants in the CS study experienced lower levels of difficulties in ER.

The second sample in this study was a clinical sample comprising participants with a diagnosis of schizophrenia. This sample had a similar percentage of males and females, with 63% never married. About 52% (n=65) had completed the higher school certificate, or TAFE certificate, or held a trade qualification. Holding an education qualification at this level in Australia is generally achieved between the age of 18 - 20years. Most people with a diagnosis of schizophrenia experience cognitive impairment making the completion of education difficult, however, the peak age of onset for a diagnosis of schizophrenia is generally between 20-28 years, therefore it is feasible someone could complete an education qualification at this level prior to receiving a diagnosis. Of the 52% very few held full-time employment (n=7). Overall, this sample had 60% unemployed. This is similar to reports in other studies where unemployment estimates of people with schizophrenia range from 70-85% (Marneros, Deister, & Rohde, 1992). However, what is unique to this sample is that 18% (n=23) were awarded a university degree, with four of these individuals holding full-time employment. It is unclear however, whether individuals completed the university degree prior to the onset and subsequent diagnosis of schizophrenia and with only four individuals in full-time employment there may be an indication of cognitive decline post education and subsequent diagnosis.

The total DERS mean score in the Schizophrenia sample (M=87.04) was higher than either the CS in this study and the mean reported in the study by Gratz and Roemer (2004) indicating participants in this clinical sample experienced higher difficulties in ER. This result might indicate that the DERS can detect differences between a clinical and CS. The significance of this will be examined further in Chapter 3.

The third sample was a clinical sample comprising of participants with a diagnosis of co-existing depression and substance abuse (CDSA). Numerous studies have reported the co-morbidity of substance abuse, with mood disorders (Ross, Glaser, & Germanson, 1988; Regier et al., 1990; Penick et al., 1994; Tomasson & Vaglum, 1995).

However, the majority of these studies only reported participants having an episode of depression in the last 12 months. This sample is unique in that every participant had a diagnosis of moderate to severe depression, as measured by the BDI-II (m=32.54), in addition to being diagnosed with substance abuse.

The total DERS mean score in the CDSA sample (M=109.9) was higher than either the CS or the schizophrenia sample, indicating participants in the CDSA sample experienced higher difficulties in ER. It is not surprising that the CDSA sample had a high DERS mean score as all the participants had moderate to severe levels of depression symptoms as scored by the BDI-II (males M=30.77, SD=8.27; females M=34.70, SD=11.00) in addition to having substance abuse. The difference in the DERS mean score between the two clinical samples might be associated with diagnosis and time of assessment. For example, while participants in both samples were living in the community, the CDSA sample was actively seeking treatment at the time of assessment, whereas participants in the schizophrenia sample were being managed in the community and not actively experiencing psychosis.

These data samples will be used in a number of ways. First, the overall DERS mean scores will be compared to establish whether the DERS can indicate differences between community and clinical samples. Second, combined data will establish the psychometric properties of the DERS. Third, the CDSA data will be used to determine whether the DERS is sensitive to change post a ten week intervention, using initial data, 3- 6- and 12- month follow-up data.

Chapter 3 will examine the reliability and validity of the DERS. Reliability will be explored in the CS to determine the consistency of the DERS across time. The best way to achieve this is by test-retest, which is when the instrument is completed on two different occasions by the same participant. Validity will also be examined by conducting a factor analysis.

CHAPTER 3 THE PSYCHOMETRIC PROPERTIES OF THE 'DIFFICULTIES IN EMOTIONAL REGULATION SCALE'

3.1 Overview

This Chapter examines the reliability and validity of the instrument referred to as the Difficulties in ER Scale (DERS) (Gratz et al., 2004) in an Australian community population. The DERS was found to have high internal reliability in the community sample (CS) (a=.94) with item-total correlations ranging from .31- .84.

The DERS test-retest reliability (n=167) was examined to determine the temporal stability of the scores and the subsequent coefficient of Alpha=.89 indicated a minimal degree of change from time one to time two. Finally, the factor structure of the DERS was explored among a diverse sample using all the DERS data collected from community and clinical samples. A four factor solution was found to be optimal. Item redundancy was explored and a shortened version of the DERS to be administered among Australian samples was suggested.

3.2 Introduction

Studies on the Difficulties in ER Scale (DERS) have shown that it has promise as a measure of difficulties in ER (Gratz et al., 2004). Chapter 2 reported on the recruitment methods, the characteristics and symptomatology of five samples: control or community sample (CS, time 1); two clinical samples, one with a diagnosis of schizophrenia and the other with a clinical diagnosis of moderate to severe depression and substance abuse (CDSA), a sample that combined CS, time 1, and both clinical groups, and a student sample. In Chapter 2 a difference in the mean scores of the DERS was reported between the samples indicating the DERS may be able to distinguish between clinical and community samples. The significance of this will be examined in this Chapter.

If clinicians are to use the DERS as an accurate measure that will assist in the planning and evaluation of therapy it is important to replicate the psychometric properties of the instrument in different samples, under different circumstances and across cultures to ensure the robustness and generalisability of the measure (Floyd et al., 1995; Haynes et al., 1995). In evaluating the psychometric properties of any instrument it is necessary to consider two aspects of information: reliability and validity. Each of these properties will now be discussed in detail.

3.2.1 Reliability

Reliability refers to the consistency of measurement and there are a number of ways to determine reliability. The first is internal consistency where Cronbach's Alpha is used to indicate the degree of relatedness among the items or questions on a scale (Streiner & Norman, 1991). The higher the alpha, or the closer it is to one, the higher the reliability estimates of the scale and so the lower the measurement error. Another concept, to evaluate reliability, is temporal stability, indicated using test-retest reliability which measures whether the instrument produces similar results when administered a second time to the same subjects. This approach assumes there is no substantial change in the construct being measured between the two occasions (Field, 2005). Test-retest reliability estimates are used to evaluate the error associated with administering a test on two different occasions to the same subjects.

In the case of the DERS, reliability matched data collected in a two week period will be examined. There is no rule in determining an appropriate time interval between the first administration of a test and the second. However, when it comes to self-rated questionnaires, generally, a retest interval of 2 - 14 days is used (Streiner et al., 1991).

Homogeneity or internal reliability of a scale examines all the items to ensure they tap different aspects of the same attribute. Items are expected to at least moderately correlate with each other and with the total scale score. To check the internal reliability of the scale, item-total correlations are used. That refers to the correlation between the individual item and the scale total omitting that item. It is generally reported that an item should correlate with the total score above r=.20 and any item with lower correlation should be discarded as it is likely to perform poorly in a factor analysis (Kline, 1994). Cronbach's alpha is a coefficient used to check a scale's internal reliability or homogeneity. If a test has strong internal consistency it is expected to yield a Cronbach alpha coefficient between .70 - .90 (Pallant, 2001).

Previous studies have reported good reliability scores for the DERS on testretest occasions over a 4-8 week period (intraclass correlation coefficients of .88, (n=21) (Gratz et al., 2004). The test-retest reliability of the DERS subscales in the above study was adequate with intra-class correlation coefficients of .69 for 'nonacceptance', .69 for 'goals', .57 for 'impulse', .68 for 'awareness', and .89 for 'strategies', and .80 for 'clarity'. Another study translated the DERS for the German population and reported acceptable internal consistency (.81 - .95), reasonable stability over a two week period (.72 to .87) and acceptable internal consistency in the subscales (.76 - .87) (Ehring et al., 2008). Table 3.1 outlines a number of published studies that have reported internal consistency on the DERS. However, only three of the five studies report the internal consistency for the overall DERS score plus the subscales, and none of these studies conducted test-retests. In summary, there has not been adequate research conducted that examines the reliability of the DERS and, in particular, reliability has not been reported in an Australian population.

Study	Sample (N)	(N)	DERS	Non	Goals	Impulse	Awareness	Strategies	Clarity
			Total	acceptance					
Whiteside, Chen, et al, 2006	Students	698	0.86	0.88	0.86	0.79	0.87	0.81	0.92
Salters-Pedneault, Roemer, et al, 2006	Students	325	0.93	All subscales > 0.76					
Tull & Roemer, 2007	Students, Panic	91	*	0.73	*	0.56	*	*	0.84
	Non- Panic	91		0.90		0.47			0.74
Tull, Barrett et al, 2007	Students	108	*		All su	bscales range	e 0.79 – 0.95		
Fox, Hong et al, 2008	Alcohol dependence	50	0.80	0.86	0.84	0.80	0.86	0.69	0.85
	Social drinkers	62	0.83	0.87	0.84	0.90	0.75	0.87	0.87

Table 3-1 Internal consistency reported in published articles

* not reported

3.2.2 Validity

Validity refers to the accuracy of the scale to measure what it is designed to measure. For the DERS, validity refers to how well it measures a person's experience of difficulties in ER. Validity can be assessed in a number of ways and is generally expressed as a validity coefficient determined by the correlation coefficient between the scores of the scale and a measure of the criterion variable. Although there are differences in the terminology used to describe the components of validity, this thesis will refer to content validity, criterion-related validity and construct validity (Streiner et al., 1991) with the knowledge that construct validity is a comprehensive concept and can include both content and criterion-related validity (Anastasi, 1990).

Content validity refers to the ability of a measurement to represent all the content of a particular construct. In the case of the DERS we asked whether the items chosen are a fair representation of the content area of difficulties in ER. One way of establishing items is based on well-accepted theoretical definitions (Cronbach, 1951; Gandek & Ware Jr., 1998). The theory behind the conceptualisation of ER has been defined in detail in Chapter 1. The DERS was drawn from various studies, some of which defined ER as involving the control of experience and expression, and the reduction of emotional arousal (Kopp, 1989; Zeman et al., 1998). Other studies emphasise emotional control, suggesting the modifying of emotions and highlighting the importance of being aware of one's emotions and understanding them (Cole et al., 1994; Thompson, 1994). The dysregulation of emotions tends to be linked with the control of emotions, given that attempts to control emotional expressions may increase the risk of emotion dysregulation. In other words, high levels of arousal can make it more difficult to regulate emotions, as the angrier one gets the longer it can take to return to a calm state (Flett, Blankstein, & Obertynski, 1996; Catanzaro, 1997; Gross & Levenson, 1997). Research has also suggested that individuals who are less aware of their emotions or lack an understanding of their emotional experience are more likely to become dysregulated (Thompson, 1994; Hayes et al., 1996). Finally, in regulating emotion, emphasis has been given to the ability to inhibit inappropriate or impulsive behaviours, or control behaviour despite emotions, and behave in accordance with desired goals, despite experiencing negative emotions (Linehan, 1994). The development of the DERS was drawn from the conceptualisation of ER that is outlined above, and involved four key areas: an 'awareness and understanding of emotions', an 'acceptance of emotions', an 'ability to control impulsive behaviours and behave in accordance with desired goals when experiencing negative emotions', and an 'ability to use situationally appropriate ER strategies flexibly, and to modulate ER responses as desired, in order to meet individual goals and situational demands' (Gratz et al., 2004).

Criterion-related validity compares scores with some other variable which is called a criterion. Concurrent validity refers to data that are gathered from different instruments at the same time (or concurrently) and generally measures a similar domain (Nunnally & Bernstein, 1994).

Predictive validity is used to determine the degree to which the measure predicts a behaviour or attribute that it should predict. In Gratz & Roemer's (2004) study, predictive validity was reported on participants with borderline personality disorder. Two behavioural outcomes, namely self harm and partner abuse were used. The study showed significant correlations between overall DERS scores and self-harm among both men and women in the expected direction. The overall DERS score and intimate partner abuse was significant among men.

In some instances predictive validity can also be referred to as construct-related discriminant validity. If the purpose is to confirm that the DERS is associated with other measures for the purpose of future planning then this is often referred to as construct-related discriminant validity (Gandek & Ware Jr., 1998). Discriminant validity will be explored below.

Construct-related validity is shown by establishing that the data generated by the new measure fit existing research and theory. The following types of construct validity will be discussed: convergent, discriminant, factorial, and known groups validity.

1. **Convergent validity** evaluates the degree to which the construct of interest is related to other conceptually relevant constructs. High correlations between scores indicate evidence of convergent validity. In the development of the DERS, the Negative Mood Regulation Scale (NMRS), (Catanzaro et al., 1990) was used to

assess the construct validity of the DERS. The NMRS was found to correlated significantly with the DERS in the expected direction.

- 2. Discriminant validity examines the correlates between scales that are different from the attribute being measured in the designed scale. Expected patterns of relations can be generalised to known groups and criterion related variables (Nunnally et al., 1994). Thus, discriminant validity can be evaluated by examining the correlations between the DERS and other measures. It is expected that measures with dissimilar constructs should have low correlations, while measures with similar constructs should show higher correlations (Anastasi, 1990).
- 3. Factorial validity is related to both internal consistency and construct validity (Nunnally et al., 1994). Factorial validity can be assessed by conducting an exploratory factor analysis, such as a principal component or a confirmatory analysis (Floyd et al., 1995). The interrelations will inform whether the inclusion of the items on a scale are justified as well as the relationship of the individual item to the scale. Factor analysis involves reducing a large data set for a group of interrelated variables into a smaller set of factors. To achieve parsimony it is best to explain the maximum amount of common variance in a correlation matrix using the smallest number of explanatory concepts (Field, 2005).

Stevens (1992) recommends interpreting only factor loadings with an absolute value greater than 0.4 as this tends to explain around 16% of the variance in the variable. This is supported by Floyd and Widaman (1995).

In common factor analysis the focus is on the variances that are common among the observed variables whereas principal component analysis examines the total variance in the data set. Common factor analysis would enable the investigator to test the underlying six-factor model presented by Gratz and Roemer (2004). The first step in this thesis will be to determine whether the DERS can demonstrate sound reliability and validity using a larger and more diverse population compared to the student sample reported in Gratz and Roemer (2004).

Originally the DERS was designed to reflect four dimensions of ER. However, the DERS was found to comprise of six components which were labelled: 'nonacceptance' of emotion responses; 'goals', indicating difficulties engaging in goal-

61

directed behaviour; 'impulse', indicating difficulties remaining in control of one's behaviour when experiencing negative emotions; 'awareness', reflecting the ability to attend to and acknowledge emotions; 'strategies', reflecting the belief that there is little that can be done to regulate emotions effectively once an individual is upset; and 'clarity', indicating an individual's ability to know the emotion they are experiencing (Gratz et al., 2004).

4. Known groups validity provides a useful parallel of information for interpreting the meaning of scale scores. This is where data are collected from two or more different samples or groups, and mean scores are compared to determine whether the scale shows the expected difference between the groups.

It is important, especially for cross-cultural purposes, to be able to define groups according to criteria that are measurable in similar ways. This allows researchers to assess the similarities and differences between groups. It also assists in identifying potential translational differences. Gratz and Roemer (2004) identified several limitations in their study that have not been addressed in any published study to date. These limitations will be addressed in this Chapter. First, test-retest reliability was based on a small student sample size (n=21). This study will draw from a larger more diverse population of 198 participants, with retest data from 169 participants. Second, validity of the DERS for a clinical population needed to be determined. As previously stated, this Chapter will address this limitation using data from two clinical groups namely, participants with a diagnosis of schizophrenia, and a second group of participants with a diagnosis of co-existing depression and substance abuse (CDSA).

Floyd and Widaman (1995) have suggested that sample limitations might affect the generalisability of results to either normal or clinical populations. It is also common knowledge that group differences often fail to replicate across groups with known differences, such as clinical and community (Floyd et al., 1995). To address this limitation data from the CS and the two clinical samples (schizophrenia and CDSA) were combined and validity was examined using factor analysis.

3.3 Aims of Chapter Three

In order to examine the psychometric properties of the DERS this Chapter had the following aims:

- 1. To compare the data from the CS in this study with the mean and standard deviations reported in Gratz and Roemer (2004) and to further examine the impact of age, gender and education in the CS on the overall mean score.
- 2. To determine the internal reliability of the scale by examining the Cronbach's alpha coefficient and the temporal reliability by examining the test-retest scores in the CS.
- 3. To determine if the DERS can detect a difference between the CS and clinical samples.
- 4. To examine the factor structure of the DERS using both the clinical and CS and explore item redundancy in order to achieve parsimony of the DERS.

3.4 Methods

Methods for this thesis have been outlined in detail in Chapter 2. A summary of the samples employed in this Chapter follows.

3.4.1 Participants

Participants were drawn from three main groups or samples: the first were volunteers from the community which comprised of university students plus residents from local registered clubs and businesses (n=198); the second sample had a diagnosis of schizophrenia and was accessed from the Schizophrenia Research Institute database (n= 129); the third sample were sourced from an existing larger study and had a diagnosis of CDSA (n=231). In order to determine reliability, the community participants were also asked to complete a second set of questionnaires 7-14 days after completing the first.

If the DERS is valid it is expected that the scores will differ according to group membership in a predictable way. In this study it is predicted that the community sample will have lower mean scores on the DERS compared to the clinical samples. As presented in Chapter 1 there is a great deal of research to state that both these clinical samples experience difficulties in ER. Therefore, it is expected that they will be representative of a clinical population that would score high on the DERS. The characteristics of each sample have been outlined in Chapter 2.

3.4.2 Statistical analysis

Temporal reliability will be examined drawing on the control sample's testretest data and conducting correlations and Pearson's r analysis. Generally when there are just two ratings, Inter-Class Correlation (ICC) is preferred over Pearson's r. However ICC is used when sample size is small (< 15) which is not the case in this study. Research has also indicated that the values of the Pearson r, and ICC are usually similar for the same set of data (Anastasi, 1990). Reliability and factor analysis will be examined on the combined samples.

3.5 Results

3.5.1 A comparison of the DERS psychometric properties reported in the study conducted by Gratz and Roemer (2004) with community control sample (CS) from this study

The DERS mean scores from the CS, Schizophrenia and CDSA samples were entered into a histogram to determine normality of distribution. Sample data that are normally distributed reduce the risk of type 1 and 2 errors occurring. Score distribution for these samples was approximately normal.

In order to establish that the CS scores in this study reflect wider population scores when compared with four other published studies, a comparison of the mean scores and standard deviations were conducted using t-tests. Table 3-2 also shows that the DERS mean of this study was lower than any other reported mean published on a student sample. The Gratz and Roemer study (2004) did not report an overall DERS mean for the entire sample, however, this was calculated using the mean and number of males and females reported in the study [(mean x n=260) + (mean x n=97) ÷ 357] giving an overall mean for the sample of 78.71 (refer to Table 3-2).

Study	n	Mean	SD	One-sample t-test
Community sample (this thesis)	198	66.35	18.22	
Gratz & Roemer (2004)	357	78.71	**_	t (198)=-9.54, *
Salters-Pedneault, et al, (2006)	325	82.66	23.41	t (197)=-12.59, *
Whiteside, et al (2006)	695	81.52	20.34	t (197)=-11.71, *
Tull, Barrett, et al (2007)	108	87.50	26.54	t (198)=-16.33, *

 Table 3-2 One sample t-test of overall DERS published mean scores and community sample in this thesis.

* indicates significant level p <.001

** not reported

To compare the total DERS mean score for the CS (M=65.35) and the mean score (M=78.71) reported in Gratz and Roemer's study (2004), a one-sample t-test was conducted. Table 3-2 reports there was a statistically significant difference between the two mean scores with participants from the Gratz and Roemer study (2004) experiencing a higher degree of difficulty in ER. Further t-tests comparing the DERS mean in this study (CS) with the mean reported in other published studies found statistical significance, with the CS sample experiencing less difficulties in ER.

3.5.1.1 Community sample total DERS and the influence of gender, age, and education

In order to determine if the difference found in the mean of the CS in this study was influenced by gender, age or education, a t-test on gender, and a series of one-way analyses of variance (ANOVA) were conducted to examine if the independent variable of the DERS total mean and the 'between groups' mean scores among the different age and education groups of the CS.

An independent t-test found no statistical difference in the overall DERS scores for males (M=67.76, SD=19.99) and for females (M=66.52, SD=18.37) in the CS study (t₍₁₉₆₎₌0.83). Participants in the CS were divided into three age groups: group 1 (18-30yrs), group 2 (31-40yrs) and group 3 (> 40 yrs). ANOVAs were conducted with significance reported at the 5% level. The DERS total scores for the three age groups were significantly different [<u>F</u> (2, 193)=4.23, <u>p</u> <.02]. Post hoc comparison using Tukey HSD test indicated the mean score for age group 1 (M=70.77, SD = 17.81) was significantly higher than that of age group 3 (M=61.94, SD=15.22). Age group 2 (M= 66.48, SD=20.86) did not differ significantly from either groups 1 or 3.
Next, education groups were entered into an ANOVA to examine the independent variable of DERS total mean and the 'between groups' mean scores among the different education groups. Participants were divided into four education groups: group 1 attended education between year 8-11, with no School Certificate completion; group 2 completed the High School Certificate (HSC); group 3 were currently attending university and group 4 had been awarded a university degree. There was a significant difference in the DERS total scores for the four education groups [<u>F</u> (3, 194)=4.74, <u>p</u> <.003]. Post hoc comparison using the Tukey HSD test indicated that the mean score for education group 3 (M=73.17, SD=20.98) was significantly higher than that of group 4 (M=62.61, SD=15.93). Education group 1 (M=62.33, SD=15.59) and group 2 (M= 65.23, SD=16.20) did not differ significantly from either groups 3 or 4.

3.5.2 Reliability of the DERS in an Australian community sample

Cronbach's alpha was calculated on the CS sample at baseline to determine the internal consistency of the DERS items. Results indicated the DERS had high internal consistency (α = .94). Item-total correlations ranged from *r*=.31 to *r*=.84. This result is consistent with the results obtained in the study by Gratz and Roemer (2004). Cronbach's alpha for each subscale ranged between .76 and .90 indicating high internal consistency. Table 3-3 reports the reliability coefficients of each subscale.

Subscale	No. of items	Cronbach's alpha	Range of item- total correlations	Range of inter- item correlations	Mean inter- item correlation
Nonacceptance	6	.87	0.57-0.82	0.45-0.71	0.54
Goals	5	.90	0.65-0.84	0.52-0.79	0.65
Impulse	6	.82	0.44-0.70	0.25-0.66	0.45
Awareness	6	.76	0.31-0.64	0.24-0.68	0.36
Strategies	8	.86	0.49-0.71	0.30-0.57	0.43
Clarity	5	.79	0.46-0.69	0.31-0.52	0.43

Table 3-3 Internal consistency reliability analyses for DERS subscales on CS (n=198) at baseline

3.5.3 Test-retest reliability

The DERS was administered on two different occasions to the CS. Internal consistency for DERS scores taken at Time 1 (α = .94) and Time 2 (α =.96) are consistent with one another. The test-retest reliability coefficient was moderate at .63.

A scatter plot of Time 1 and Time 2 scores shows the strong linearity in the data (Figure 3:1). A paired sample t-test was conducted to evaluate the group average of the DERS total scores at Time 1 and Time 2. There was no statistically significant difference from Time 1 (M=65.26, SD=17.23) to Time 2 (M=64.19, SD=18.96), t (166) =1.60, p < .110, suggesting minimal degree of change in response to items from Time 1 to Time 2.



Figure 3:1 Scatter plot of time one and time two.

Table 3-4 reports on the correlations for the each of the DERS six subscales at Time 1 and Time 2. All items correlated at least .3 suggesting response to items on each subscale taken at different times shared common variance.

Table 3-4 Correlations: CS at time 1 (baseline) with time 2 (retest)

	Total	Nonaccept- ance	Goal	Impuls- ive	Strategies	Awareness	Clarity
Total	.89**	.72**	.69**	.74**	.79**	.62**	.73**
Nonacceptance	.69**	.75**	.46**	.58**	.69**	.41**	.56**
Goal	.61**	.46**	.79**	.48**	.49**	.25**	.48**
Impulsive	.70**	.56**	.59**	.72**	.62**	.38**	.54**
Strategies	.81**	.62**	.57**	.71**	.86**	.53**	.66**
Awareness	.57**	.38**	.30**	.39**	.45**	.81**	.42**
Clarity	.79**	.54**	.42**	.56**	.60**	.56**	.75**

** Correlation is significant at the 0.01 level (2-tailed).

Time 1 - down

Time 2 – across

A paired t-test was conducted to evaluate CS on Time 1 with Time 2 on the DERS subscales of 'nonacceptance', 'goals', 'impulse', 'strategy', 'awareness' and 'clarity' scores. There was a significant difference between Time 1 (M=12.38, SD=4.41) and Time 2 (M=11.25, SD=4.24) on the subscale of 'goals', t (166) =5.22, p < .000. Participants at Time 1 experienced more difficulties engaging in goal directed behaviour compared to Time 2.

3.5.4 Can the DERS detect differences between the control and clinical samples?

In order to establish whether the DERS can detect differences between clinical and community samples, the mean scores of each sample were compared using an independent t-test. The results are presented in Table 3-5 and show that all t-tests were statistically significant, indicating that the data from the clinical samples were significantly different from the CS, with the clinical samples experiencing higher difficulties in ER. The highest mean was found in the CDSA sample.

Study	Mean	SD	Independent t-test
Control baseline	66.35	18.22	
Schizophrenia	87.04	23.94	t(321) =-8.27, p <.001*
CDSA	108.99	21.46	t(428) =-22.28, p <.001*

Table 3-5 Independent t-test comparing control sample with clinical samples

3.5.4.1 Reliability using all data from community and clinical sample

As mentioned above, the rationale for combining the samples was twofold: First, the clinical sample was living in the community at the time of assessment and combining them with the CS would more accurately reflect a sample that is more representative of the general population. Second, it allowed a more parsimonious factor solution.

Reliability analysis on all items in the combined sample appeared acceptable with item-total correlations ranging from .66 - .78, α =.96. The item with the lowest value was item 17 (.30). However, removal of this item would not alter the overall alpha (a=.96).

Reliability analysis in the combined sample on each subscale was conducted. All items on the six subscales appeared acceptable: 'nonacceptance' (α =.91), 'goals' (α =.86),' impulse' (α =.89), 'strategies' (α =.92),' awareness' (α =.83) and 'clarity' (α =.86). The greatest increase in Cronbach's alpha would come from deleting item 20 on the goal subscale as it scored the lowest correlation (.30). However, removal would increase the overall alpha by only .01. The combined sample scores on the DERS had good reliability with all subscales within acceptable range. See Table 3-6 for internal consistency for each subscale.

Subscale	No. of items	Cronbach's alpha	Range of item- total correlations	Range of inter-item correlations	Mean inter- item correlation
Nonacceptance	6	0.91	0.81-0.64	0.72-0.55	0.61
Goals	5	0.86	0.52-0.75	0.69-0.43	0.55
Impulse	6	0.89	0.61-0.78	0.45-0.71	0.61
Awareness	6	0.83	0.48-0.70	0.35-0.68	0.45
Strategies	8	0.92	0.47-0.66	0.66-0.78	0.59
Clarity	5	0.86	0.66-0.72	0.52-0.66	0.56

Table 3-6 Internal consistency reliability analyses for DERS subscales in combined sample (n=555)

Reliability for the combined sample showed a similar albeit slightly increased Cronbach's alpha compared to that found in the CS.

3.5.5 Factor Analysis

3.5.5.1 A comparison of the six components from Gratz & Roemer (2004) study and the combined sample in this study

All items were entered into a factor analysis according to the subscales in the original study to determine if similar findings were apparent. Entering items according to the subscales outlined by Gratz and Roemer (2004) allowed for easy visual interpretation of the items. For example, the expectation was that all the items for the subscale of 'nonacceptance' would load together with an item loading greater than .40. If those items did not load as a group, or items cross-loaded on more than one factor, or had a low item loading, then that item was considered weak.

Factor analyse on the combined sample (n=557) resulted in five components with eigenvalues greater than 1.0. Eigenvalues greater than 1.0 indicate that the components explains more variability than one of the original items (Bryman & Cramer, 2001). The first component explained 43.13% of the overall variance.

A scree test is generally thought of as a more accurate method for retaining factors than the more commonly used criterion of eigenvalues greater than 1.0 (Nunnally et al., 1994). The scree plot shown in Figure 3:2 indicates retaining two components given the elbow at component 3, however, there were five eigenvalues greater than 1.0 explaining 66.55% of the variance.

Scree Plot



Figure 3:2 Scree plot using combined data sample

A decision was made to conduct factor analysis on one to six components to examine interpretability with the six component solution reported in Table 3-7. This Table compares the results of the six component solution on the 36 items reported in the DERS (2004) with the solution found using the combined community and clinical sample in this study. For clarity only loadings greater than .30 are shown. The item order shown at the left-hand side of Table 3-7 is that reported by Gratz and Roemer (2004) and shows the six components referred to as 'nonacceptance', 'goals', 'impulse' 'awareness', 'strategies' and 'clarity' (refer to the definition for each factor outlined earlier in the Chapter). The right-hand side reports the results of a six component solution found in this study. The results revealed that items did not load on the same factors when compared to the original study by Gratz and Roemer (2004). Component one held the majority of items from the subscales of 'impulse' and 'strategies' with the exception of item 30 and item 22 that loaded elsewhere. Component two held all the items from the 'awareness' subscale plus 3 items from 'clarity', with item 4 crossloading. Component three held all the items from the 'nonacceptance' subscale plus item 30. Component four held all the items from the 'goals' subscale and component five held 3 items from 'clarity' subscale however item 4 crossloaded. And finally, component six only held 1 item (22) at > .40.

	Published *					Combined sample							
Subscale	ITEM	1	2	3	4	5	6	1	2	3	4	5	6
Non	25	0.91								0.85			
acceptance	21	0.76								0.88			
	12	0.72								0.79			
	11	0.53								0.69			
	29	0.48				0.35				0.69			
	23	0.41								0.58			
Goals	26		0.88								0.88		
	18		0.88								0.80		
	13		0.85								0.83		
	33		0.65					0.31			0.53		
	20		0.64								0.73		
Impulse	32			1.00				0.76					
	27			0.79				0.68					
	14			0.75				0.73					0.32
	19			0.58				0.60					
	3			0.53				0.54					
	24			0.40				0.41				0.32	
Awareness	6				0.74				0.83				
	2				0.67				0.83				
	10				0.61				0.66				
	17				0.59				0.52	0.33			
	8				0.58				0.79				
	34				0.57				0.46				-0.32
Strategies	16					0.86		0.55					
	15					0.79		0.70					
	31					0.64		0.63					
	35					0.61		0.61					
	28					0.59		0.68					
	22				0.43	0.49		0.39					-0.44
	36					0.45		0.44					
	30	0.34				0.45				0.46			
Clarity	5						0.81		0.31			-0.56	
	4						0.71		0.48			-0.43	
	9						0.69					-0.52	
	7				0.32		0.59		0.68				
	1				0.38		0.42		0.57				

Table 3-7 Comparison of six component solution

* refers to the original study by Gratz and Roemer (2004)

The results indicate that an attempt to replicate the same six component solution reported in the Gratz and Roemer study (2004) was not reached. In general, it appeared that the items for 'clarity' did not load together on one component and a number of these items cross-loaded over two components, therefore any further analysis needs to be conducted minus these items. It also appeared that a five or four component solution would be optimal, however, if we remain with the existing subscales, all clarity items would need to be removed from further analysis.

3.5.6 Alternative solution on the combined sample

Field (2005) suggests that items are expected to load on the one factor or component. If items load on more than one factor it is likely to be measuring something other than what was intended. Hence, prior to deleting items to determine the best possible solution for factor structure, it was decided to ignore the six subscales of the DERS as reported by Gratz and Roemer (2004) and re-examine the 36 items. Oblique rotation was used to allow for the possibility that factors may be correlated with each other. Field (2005) suggests the use of factor rotation which can discriminate between components and ensure items load maximally on only one component.

The component structure was explored using a principal components analysis with an oblimin (oblique) rotation. Five eigenvalues greater than one were found to explain 63.62% of the variance and showed all items loaded >.40. However, component five contained only 3 items. A four component solution showed 5 items loading on component four, 7 items on component three, 9 factors on component two, and 14 items on component one. Item 33 crossloaded and item 24 loaded <.30.

Table 3-8 reports a four component solution, minus items 24 and item 33 as they did not contribute to a simple factor structure and failed to meet a minimum criterion of having primary factor loadings > .40 without crossloading on another factor. All items loaded onto four components with a value of .40 or greater, and no items crossloaded.

	Four component analysis in Combined Sample							
Item		1	2	3	4			
14	When I'm upset, I become out of control.	0.76						
32	When I'm upset, I lose control over my behaviour.	0.75						
35	When I'm upset, it takes me a long time to feel better.	0.70						
15	When I'm upset, I believe I will remain that way for a long time.	0.70						
28	When I'm upset, I believe there is nothing I can do to make myself feel better.	0.67						
3	I experience my emotions as overwhelming and out of control.	0.67						
19	When I'm upset, I feel out of control.	0.64						
31	When I'm upset, I believe that wallowing is all I can do.	0.63						
27	When I'm upset, I have difficulty controlling my behaviours.	0.61						
16	When I'm upset, I believe I will end up feeling depressed.	0.59						
9	I am confused about how I feel.	0.54						
36	When I'm upset, my emotions feel overwhelming.	0.53						
5	I have difficulty making sense out of my feelings.	0.51						
4	I have no idea how I am feeling.	0.49						
2	I pay attention to how I feel.		0.76					
8	I care about what I am feeling.		0.75					
6	I am attentive to my feelings.		0.74					
10	When I'm upset, I acknowledge my emotions.		0.73					
17	When I'm upset, I believe that my feelings are valid and important.		0.65					
7	I know exactly how I am feeling.		0.64					
1	I am clear about my feelings.		0.58					
34	When I'm upset, it takes a long time to figure out what I'm really feeling.		0.56					
22	When I'm upset, I know that I can find a way to eventually feel better.		0.45					
21	When I'm upset, I feel ashamed with myself for feeling that way.			0.85				
25	When I'm upset, I feel guilty for feeling that way.			0.84				
12	When I'm upset, I become embarrassed for feeling that way.			0.75				
29	When I'm upset, I become irritated with myself for feeling that way.			0.68				
11	When I'm upset, I become angry with myself for feeling that way.			0.67				
23	When I'm upset, I feel like I am weak.			0.56				
30	When I'm upset, I start to feel very bad about myself.			0.45				
26	When I'm upset, I have difficulty concentrating.				0.82			
13	When I'm upset, I have difficulty getting work done.				0.76			
18	When I'm upset, I have difficulty focussing on other things.				0.74			
20	When I'm upset, I can still get things done.				0.74			

The combined sample showed the first component consisted of 14 items that indicate control or lack of control, for example, "When I'm upset I become out of control", "When I'm upset I have difficulties controlling my behaviours", "When I'm upset, I lose control over my behaviour", "When I'm upset, I believe that I will remain that way for a long time", "When I'm upset, I believe that there is nothing I can do to make myself feel better", and "When I'm upset, my emotions feel overwhelming". When compared to the published subscales, the majority of these items were from the 'impulse' or 'strategy' subscales, with three items that belonged to the 'clarity' subscale (items 4, 5, 9) and had lower loadings compared to the original study.

Component two in the combined sample consisted mostly of items that indicated awareness or recognition, for example, "I am attentive to my feelings" (item 6), "I pay attention to how I feel" (item 2), "When I'm upset, I acknowledge my emotions" (item 10), "I care about what I am feeling" (item 8), "When I'm upset, I take time to figure out what I'm really feeling" (item 34), and "When I'm upset, I believe that my feelings are valid and important" (item 17). Two items that would have previously been referred to as 'clarity' items held similar semantics to the other items on this factor, "I know exactly how I feel" (item 7), "I am clear about my feelings" (item 1) and item number 22 "When I'm upset, I know that I can find a way to eventually feel better" which might explain why they loaded together on component two.

All items that loaded on component three were the same as the items from the subscale of 'nonacceptance' in the original scale with the exception of item 30. These items semantically indicated that in the presence of emotion (upset), individuals judge the emotion as negative and place a value judgement on being upset, for example, "When I'm upset, I feel guilty for feeling that way" (item 25), "When I'm upset, I feel ashamed with myself for feeling that way" (item 21), and "When upset, I become embarrassed for feeling that way" (item 15), "When I'm upset, I start to feel very bad about myself" (item 30). The original study referred to item 30 as a 'strategy' however it is semantically similar to an item in the subscale of 'nonacceptance'. In other words, individuals could easily read item 30 'feeling bad about myself' as a judgement similar to item 21 'feel ashamed', or item 12 'embarrassed'. The remaining four items that loaded on component four were referred to as 'goal' items in the original research. Each item referred to "when upset I have difficulty - 'concentrating', 'working', 'focussing', or 'getting things done'.

3.5.6.1 Item redundancy

As described above, parsimony is achieved if the maximum amount of common variance in a correlation matrix is explained by using the smallest number of explanatory concepts (Field, 2005). While the four component solution seems optimal, some items appear to be asking the same thing. If this is the case then one of those items may be semantically and psychometrically redundant. Semantic redundancy is a concern of content validity and would be determined by looking at the semantics of the item to determine if that item is capturing the construct. Semantic redundancy is very subjective. However redundancy can be supported statistically by examining the correlations between the two items that appear to be asking the same thing. If the two items have a high correlation then one can be removed (Streiner et al., 1991). Moderate correlations (r=.30 - .70) for each item are sought with an overall a > .80 (Nunnally et al., 1994). Table 3-9 provides the item-total correlations of all items (minus item 24 and 33) and shows an overall a .96, with individual item-total correlations ranging from .43 to .80.

Table 3-9 Item	Correlation	on the	DERS
----------------	-------------	--------	------

InterviseVarianceVarianceIntervise1I am clear about feelings76.03669.900.682I pay attention to feelings76.23681.560.523I experience my emotions as overwhelming and out of control76.48669.790.714I have no idea how I am feeling76.81680.440.635I have difficulty making sense out of my feelings76.39671.900.696I am attentive to my feelings76.02677.620.537I know exactly how I am feeling75.86671.460.64	<u></u>
11 am clear about reelings76.03609.900.082I pay attention to feelings76.23681.560.523I experience my emotions as overwhelming and out of control76.48669.790.714I have no idea how I am feeling76.81680.440.635I have difficulty making sense out of my feelings76.39671.900.696I am attentive to my feelings76.02677.620.537I know exactly how I am feeling75.86671.460.64	
2I pay attention to reelings76.23661.560.523I experience my emotions as overwhelming and out of control76.48669.790.714I have no idea how I am feeling76.81680.440.635I have difficulty making sense out of my feelings76.39671.900.696I am attentive to my feelings76.02677.620.537I know exactly how I am feeling75.86671.460.64	
31 experience my emotions as overwhelming and out of control76.48669.790.714I have no idea how I am feeling76.81680.440.635I have difficulty making sense out of my feelings76.39671.900.696I am attentive to my feelings76.02677.620.537I know exactly how I am feeling75.86671.460.64	
4I have no idea how I am feeling76.81680.440.635I have difficulty making sense out of my feelings76.39671.900.696I am attentive to my feelings76.02677.620.537I know exactly how I am feeling75.86671.460.64	
5I have difficulty making sense out of my feelings76.39671.900.696I am attentive to my feelings76.02677.620.537I know exactly how I am feeling75.86671.460.64	
6 I am attentive to my feelings 76.02 677.62 0.53 7 I know exactly how I am feeling 75.86 671.46 0.64	
7 I know exactly how I am feeling 75.86 671.46 0.64	
8 I care about what I am feeling 76.46 683.75 0.48	
9 I am confused about how I feel 76.32 669.15 0.68	
10 When I'm upset, I acknowledge my emotions 76.09 683.81 0.43	
when the upset. I become angry with mysell for reeling that 76.18 669.01 0.66	
12 When I'm upset I become embarrassed for feeling that way 76.10 607.01 0.60	
12 When I'm upset, I become embanassed for reening that way 76.27 670.37 0.60	
13 When I'm upset I become out of control 76.02 010.04 0.02	
14 When I'm upset L believe Luill remein that you for a long time 76.65 666.02 0.70	
15 When I'm upset, I believe I will remain that way for a fong time 70.05 000.05 0.74	
When I'm upset. I believe that my feeling are valid and	
17 important 76.07 687.85 0.37	
18 When I'm upset. I have difficulty focussing on other things 75.52 670.22 0.63	
19 When I'm upset, I feel out of control 76.58 665.90 0.73	
20 When I'm upset. I can still get things done 75.61 673.43 0.59	
When I'm upset, I feel ashamed with myself for feeling that	
21 way 76.47 672.79 0.64	
When I'm upset, I know that I can find a way to eventually feel	
22 better 76.01 662.08 0.72	
23 When I'm upset, I feel like I am weak 76.21 666.72 0.66	
25 When I'm upset, I feel guilty for feeling that way 76.45 671.66 0.65	
26When I'm upset, I have difficulty concentrating75.62672.630.60	
27 When I'm upset, I have difficulty controlling my behaviours 76.48 666.86 0.71	
When I'm upset, I believe there is nothing I can do to make	
28 myself feel better 76.53 667.13 0.72	
When I'm upset, I become irritated with myself for feeling that	
29 Way 70.14 000.32 0.00	
30 When the upset I believe that wellowing is all leap do 76.10 033.00 0.77	
31 When the upset, I believe that wallowing is all I can do 70.36 670.97 0.07 22 When the upset, I believe that wallowing is all I can do 74.72 640.71 0.72	
32 when im upset, i lose control over my benaviour /0.72 006.71 0.72	
when init upset, it takes a long time to rigure out what Fam 34 really feeling 75.58 680.17 0.45	
35 When I'm unset it take me a long time to feel better 76.10 664.58 0.73	
35When I'm upset, in take the a long time to reel beller75.17004.000.7336When I'm upset, my emotions feel overwhelming75.85659.270.72	

Another way to reduce items would be to remove misfitting items from scales, in this case the items that belong to the 'clarity' subscale. 'Clarity' items did not load on their own, and tended to cross-load over a number of components. A component analysis was recomputed minus the 'clarity' items to determine the stability of the remaining 29 items (refer to Table 3.10). Component analysis of the 29 items resulted in four components with eigenvalues greater than 1.0 that accounted for 66.5% of the total

variance. Component one explained 46.1%, factor two explained a further 9.61%, factor three explained a further 5.10% and component four explained 4.36%.

COMPONENT	ITEM	1	2	3	4
Strategies	32	0.82			
(lack of access to ER strategies – 11 items)	14	0.81			
	15	0.73			
	28	0.71			
	27	0.71			
	19	0.69			
	31	0.67			
	35	0.67			
	3	0.63			
	16	0.61			
	36	0.53			
Awareness	2		0.77		
(lack of emotion awareness – 7 items)	8		0.76		
	10		0.75		
	6		0.74		
	17		0.67		
	34		0.62		
	22		0.44		
Nonacceptance	21			0.84	
(nonacceptance of emotion	25			0.82	
responses – 7 items)	12			0.75	
	29			0.67	
	11			0.67	
	23			0.55	
	30			0.42	
Goals	26				0.84
(difficulties engaging in goal-directed behaviour	13				0.83
– 4 items)	18				0.79
	20				0.75

Table 3-10 Final component analysis

(Minus items 24, 33, 1, 4, 5, 7, & 9)

Next, to examine whether the four subscales of ER related to one another, as expected, each was entered into a correlation. Table 3-11 shows each component related to the other at the 0.01 level.

	Strategies	Awareness	Nonacceptance	Goals
Strategies				
Awareness	.314(**)			
Nonacceptance	.575(**)			
Goals	.575(**)	.265(**)	.464(**)	

Table 3-11 Correlations between the four subscales

** Correlation is significant at the 0.01 level (2-tailed)

Reliability analysis was conducted entering the items from each of the four components. Component one (11 items) had an overall α .94. Reliability of items on component one indicated all items appeared acceptable with an item-total correlation range from r = .60 to r= .80. Component two (seven items) had an overall α =.86 with item-total correlation range from r = .48 to r= .71. Component three (seven items) had an overall α =.92 with item-total correlation range from r = .66 to r= .80. Component four (four items) had an overall α .86, with item-total correlation range from .51- .74. Item-total correlations indicated that the removal of any item from any of the four components would not alter the reliability of the scale as a whole.

In summary, the final four component structure (minus items 33, 24, 5, 4, 9, 7, and 1) did not compromise the reliability of the scale. The DERS had high internal consistency *a* .96. All 29 of the items had item-total correlations above .58. Each of the four final component had adequate internal consistency, with an overall *a* .85 or greater for each component.

The next step was for each component to be examined and labelled for common themes that represent 'real-world' constructs. Common themes are determined by the content of each question that loaded onto the same factor. Factor one held items from both 'impulse' and 'strategy' subscales and tended to refer to not being able to manage, or to act on a strategy, when experiencing a negative emotion; for example, 'When I'm upset, I become out of control' (impulse), and 'When I'm upset, I believe there is nothing I can do to make myself feel better' (strategy).

Component two continued to hold items from the 'awareness' subscale of the DERS. Therefore, these items continue to represent 'awareness' and scores are reversed to reflect a lack of emotion awareness, so this name was retained. All items from the 'nonacceptance' subscale of the DERS loaded on the third component. Therefore, these

items continue to represent 'nonacceptance' of emotions. Finally, all items from the 'goal' subscale of the DERS loaded on the fourth factor. Thus, the revised DERS held four subscales labelled 'strategies', 'awareness', 'nonacceptance', and 'goals'.

3.6 Discussion

This Chapter set out to address some of the limitations presented in the development of the DERS outlined by Gratz and Roemer (2004). The first aim of this study was to compare the mean scores and standard deviations of the Gratz and Roemer's study (2004) and the CS representing the Australian population. The data from the CS showed the overall mean of the DERS was significantly different when compared to the mean scores of samples from other studies. The CS participants scored lower on the DERS, indicating the CS experienced fewer difficulties in ER. This result may suggest that Australian participants experience less difficulty in ER when compared to the Boston university students in the sample used by Gratz and Roemer (2004). However, this difference may be influenced by the CS participants who were drawn from the general community that included both students and workers, as opposed to the student sample used in the original study.

The second aim of this study was to examine the reliability of the DERS using Cronbach's alpha and test-retest data to determine the consistency of the scores. One limitation outlined in the Gratz and Roemer study (2004) was that the test-retest reliability results were based on a small sample size of 21 university participants. The test-retest in this study was based on 167 community participants, and found the DERS to have acceptable temporal reliability. Internal reliability analysis of the DERS on the community test data indicated high internal consistency with *a* =.94. A reliability analysis was repeated using the combined sample and also showed a high internal consistency with *a* = .92.

The third aim of this study was to determine whether the DERS could detect a difference between the CS and the clinical samples. Independent t-tests indicated the DERS mean in the CS group was significantly different from the DERS mean in both clinical groups (schizophrenia and CDSA). The CDSA experienced greater difficulties in ER when compared to the other two samples. This was an expected outcome as participants in the CDSA sample were referred into a treatment program, held a

minimum of two clinical diagnoses, and were possibly experiencing great difficulties in ER at the time of completing the questionnaires. The schizophrenia sample, on the other hand, was being managed in the community at the time of completing the questionnaire and was not currently experiencing any psychotic symptoms. It was expected that while they may report a statistically higher mean score when compared to the community population, their score would be lower than the CDSA sample.

The fourth aim was to examine the validity of the DERS by conducting a number of factor analyses, using a diverse sample including the community sample and two clinical samples, namely, schizophrenia and CDSA. Results indicated a four component solution excluding items 24, 33, 1, 4, 5, 7, & 9 was the optimal solution.

Gratz and Roemer (2004) set out to develop a scale to reflect four components of difficulties in ER, which was designed to reflect the following; a) awareness and understanding of emotions, b) acceptance of emotions, c) ability to engage in goaldirected behaviour when experiencing negative emotions and d) ability to use situationally appropriate ER strategies (Gratz et al., 2004). However, the outcome of their study found the DERS reflected six components and suggested that 'awareness' and 'understanding' loaded on separate factors which they labelled 'awareness' and 'clarity'. 'Goal directed behaviour' and 'refrain from impulse' also loaded onto separate factors labelled 'impulse', and 'goals', whilst 'nonacceptance' and 'strategies' mapped directly onto the expected components.

While a four component solution presented in this study seems optimal it does not fit with the existing six factor solution presented by Gratz and Roemer (2004). This suggests participants were responding differently to the DERS when compared to the Gratz and Roemer (2004) study. However, with the exception of a few items, the final four component found in this study support the original conceptualisation of ER. Component one contained mainly 'impulse' and 'strategy' items (11/14 items), indicating that items may be seen as having similar meaning. In this study 'impulse' refers to difficulties in impulse control when experiencing negative emotions and 'strategies' refer to having limited access to ER strategies. When looking at the semantics of the items, 'impulse' items tend to focus on being out of control of both emotions and behaviours (impulse) for example, 'when I'm upset, I lose control over my behaviours', while 'strategies' items tend to focus on understanding emotions and how long that emotion will stay with them, for example, 'when I'm upset, I believe that wallowing in it is all I can do'. It may be that the participants in this study see being out of control similarly to experiencing emotions, suggesting 'If I'm emotional then I am not in control' and as a result have responded to the items on these two factors in a similar way.

Component two consisted of all six 'awareness' items, along with one item from strategies. This loading pattern suggests that there is no difference between being aware of emotion responses and understanding these responses. This is in contrast to Gratz and Roemer (2004) where 'awareness' and 'understanding' or 'clarity' loaded on different components. Component three consisted of all the 'nonacceptance' items, and one strategy item. Finally, component four consisted of four 'goal' items. The outcome of the final four factor analysis, in this study, tends to support the four proposed components of ER that was initially conceptualised by Gratz and Roemer (2004). These four proposed components are still captured in the DERS instrument when used in an Australian sample consisting of both community and clinical participants.

Chapter 4 will explore construct validity of this revised 29 item DERS (DERS-Revised). As described in Chapter 2, additional measures were administered to the combined community/clinical sample to examine predictive validity: the Kessler 10 (K10) and the Social Problem Solving Inventory-Revised (SPSI-R) and a fourth sample of university students were also examined using the DERS, the K10 and the IPIP.

It is expected that the DERS-R (a measure of difficulties in ER) and the K10 (a measure of psychological distress) will be related moderately with one another. It is also expected that the DERS-R will be negatively related with scores on SPSI-R (a measure of social problem solving ability), as someone who has high scores on the DERS-R (more difficulties in ER) will have low scores on the SPSI-R (less ability to problem solve). These, and other measures, will be used to examine construct validity in Chapter 4.

CHAPTER 4 THE CONSTRUCT VALIDITY OF THE DERS-R

4.1 Overview

Confirming a person's subjective experience by a scale designed to measure difficulties in ER is an ongoing process. However, comparisons with existing instruments that are known to reflect a relationship with the construct, in this case difficulties in ER, will provide further evidence of the validation of the Difficulties in ER- Revised (DERS-R) measure.

This Chapter explores the construct validity of the DERS-R in two samples. The first sample consists of data combined from the clinical and control samples employed in Chapter 3. Two instruments are used to assess the construct validity of the DERS-R. The first scale is the Kessler 10 (K10) (G Andrews et al., 2001), a scale of psychological distress, depression and anxiety, and it is expected to be positively related to high scores on the DERS-R. The second scale is the Social Problem Solving Inventory-Revised (SPSI-R), (Thomas J. D'Zurilla et al., 1998), which relates to problem solving ability and is expected to be negatively related to the DERS-R. The second scale of university students who were administered the DERS-R. The second sample consists of university Item Pool (IPIP) (Goldberg, 1992) which is a measure of the Five Factor Model of personality.

Construct validity was supported by significantly high correlation between the DERS-R overall score, and each of the DERS-R sub-factors, and the K10 in the expected direction. When the analysis was repeated in sample 2 the results were similar. There was a significant correlation between the DERS-R and SPSI-R in sample 1 in the expected direction. Finally, the correlations between the DERS-R and the IPIP revealed a significantly high positive correlation on the neuroticism subscale. In conclusion, these findings indicate the DERS-R had sound construct validity and is a suitable scale to use in a clinical sample.

4.2 Introduction

Construct validity of the DERS-R is examined in this Chapter by determining its relationship with other measures that are already established. If there is no relation with existing measures, especially when theory indicates a likely relationship, then the DERS-R will have limited clinical relevance (Anastasi, 1990).

The impact of experiencing a dysregulation of emotions and the ability to function in everyday life is widely known (Thompson, 1994). Often, the more an individual becomes emotionally dysregulated, the more they become increasingly unable to think clearly, or to address everyday problems. Research has indicated that difficulties in ER are linked with psychopathology (Bushman et al., 2001; Tull, Gratz, Salters, & Roemer, 2004; Briere, 2006; Salters-Pedneault et al., 2006), physical ailments (Gross & Levenson, 1997), chronic worry (Salters-Pedneault et al., 2006), and reduced clarity or the ability to think clearly when making decisions or problem solving (Rude & McCarthy, 2003).

4.2.1 Social problem solving, mood instability, and dysregulation of emotions

Problem solving can be conceptualised as an important coping strategy which enables a person to better manage daily problematic situations and their emotional effects, thereby reducing, or preventing, psychological distress (D'Zurilla & Sheedy, 1991). One of the two orientations of the SPSI-R measure is referred to as 'positive problem orientation' and can be described as constructive, problem-solving cognitions which include generalised challenge appraisals, self-efficacy, and positive outcome expectancies. The second orientation is referred to as 'negative problem orientation' and refers to dysfunctional cognitive-emotional schemas that include generalised threat appraisals, low self-efficacy, negative outcome expectancies and low frustration tolerance (Thomas J. D'Zurilla et al., 1998). Deficits in problem solving have been reported in both depressed and anxious individuals when comparing normal and clinical control subjects (Marx, Williams, & Claridge, 1992). Another study that examined the relationship between everyday problems, social problem solving, and depression and anxiety in middle-aged and elderly community residents found that a 'negative problem' orientation accounted for a significant degree of causal relations between everyday problems and depression and anxiety (Kant, D'Zurilla, & Maydeu-Olivares, 1997). These studies suggest the more psychologically distressed an individual is the greater difficulties they would experience in social-problem solving.

In the present study it is therefore hypothesised that high scores on the K10, indicating greater psychological distress will be associated with high scores on the 'negative orientation' of the SPSI-R scale.

Research has also indicated that effective problem solving can reduce stress and negative stress outcomes. This is because effective problem solving is viewed as an important coping strategy that increases general competence and adaptation (Nezu & Carnevale, 1987; Nezu & Perri, 1989; Thomas J. D'Zurilla et al., 1998). If effective problem solving is linked to reduced stress then it can be hypothesised that a negative relationship will exist between psychological distress, as measured by the K10, and effective problem solving as measured on SPSI-R.

The research outlined above has indicated that psychological distress may be related to poor social problem-solving. It would also make sense that a person who is psychologically distressed would also experience difficulties regulating their emotions, as indicated with high scores on the DERS-R. The link between experiencing difficulties in ER and psychopathology has been outlined in Chapter 1. However the findings of two relevant studies will be briefly summarised here. Salters-Pedneault and Roemer (2006) examined the relationship between chronic worry, generalised anxiety disorder, and difficulties in ER in a non-clinical sample and reported a significant association, indicating those who experience higher levels of chronic worry would gain higher scores on the DERS. Another study examined non-clinical individuals who exhibited high levels of post-stress disorder (PSD) symptoms that were consistent with a post-traumatic stress disorder diagnosis, and reported significantly higher scores on the DERS compared to individuals without PSD symptoms (Tull, Barrett et al., 2007).

There is also evidence to suggest that depressed individuals show greater dysfunctional use of ER strategies, with some researchers suggesting that depressive episodes are a consequence of maladaptive ER (Gross & Monoz, 1995; Campbell-Sills, Barlow, Brown, & Hofmann, 2006; Kring & Sloan, 2007). A study comparing individuals with a past history of depression with controls, reported that depressed individuals scored significantly higher on the DERS compared to controls (Ehring et al., 2008). This link between anxiety and depressive disorder, and difficulties in ER suggests that high scores on the K10 would be positively related to high scores on the

85

DERS-R presenting additional evidence to support the construct validity of the DERS-R.

4.2.2 Big Five from the IPIP (Goldberg, 1992)

The Big Five personality factors and their characteristic traits are described as 'neuroticism' which refers to individuals who are more worried, insecure, nervous, and highly strung; 'extroversion' which refers to individuals with the characteristics of sociable, talkative, fun-loving and affectionate; 'openness', described as original, independent, creative, and daring behaviour; 'agreeableness', indicated by goodnatured, soft-hearted, trusting and courteous behaviour; and 'conscientiousness' indicated by the characteristics of careful, reliable, hardworking, and organised stability (Costa & McCrae, 1984; Watson, Clark, McIntyre, & Hamaker, 1992). Numerous studies have reported correlational findings between negative ER and personality differences. One study reported that neuroticism predicted lower use of strategies to repair negative emotions and neuroticism was associated with maladaptive ER strategy use (Kokkonen & Pulkkinen, 2001). Also correlating negatively with neuroticism is the tendency to reduce or eliminate one's negative emotions, or turn them towards a more positive direction (Davies, Stankov, & Roberts, 1998). Finally, neuroticism has been associated with an increase in negative emotions, and that increase can result from differences in the tendency to regulate emotions (Weiting & Diener, 2009). These studies suggest that individuals who experience greater difficulties in ER will also report higher neuroticism. Therefore, a positive relationship between the DERS-R and 'neuroticism' would be expected in this study.

People high in extroversion and low in neuroticism were found to be genetically predisposed to emotional stability (Costa et al., 1984; Watson et al., 1992). This indicates that emotional stability is linked with extroversion and would suggest that people who score high on extroversion would produce low scores on the DERS-R. It is therefore expected that the DERS-R will positively relate to neuroticism and negatively relate to extroversion, as measured by the Big Five questionnaire from the IPIP (Goldberg et al., 2006).

86

4.2.3 Rationale for employing two samples in this Chapter

A number of researchers have reported that recruiting university students for the purpose of validating a measure to be utilised on a clinical population may not be the most effective way to practise translational research (in this study translating basic findings from a non-clinical sample to a clinical population) (Rosenthal et al., 2008). University students may not be representative of the wider community and care should be taken when generalising the mean scores of a measure to a clinical sample. Students are generally considered more functional, and this is based on the fact that they can undertake a degree and meet the assessment requirements to complete a degree (Tull, Barrett et al., 2007). Other researchers have argued that among university students there is a high degree of mental health problems. Using the K10 and recruiting over 6000 students it was reported that mental health problems were 19.2%, with 67.4% reporting subsyndromal symptoms (Stallman, 2010). This percentage was found to be higher than that reported in a household survey where 17.7% of Australian adults (n=10,600) reported common mental disorders of anxiety, depression, alcohol or substance abuse and neurasthenia (Gavin Andrews et al., 2001). To ensure a measure more accurately reflects a sample that represents the general community this study has recruited and combined two data groups referred to as sample one. Sample 1 will consist of volunteers from the community population including recruitment from local businesses, registered clubs and university students. This sample will be combined with individuals who have a clinical diagnosis of schizophrenia and were recruited from the Schizophrenia Research Registry. The clinical group was living in the community at the time of recruitment, and considered part of the general community population. Sample 2 was composed of university students (demographic characteristics of this sample are outlined in Chapter 2). Using these two samples will also address one of the recommendations from the Gratz and Roemer (2004) study regarding replication of the psychometric results with different samples under different circumstances.

4.3 Aim of Chapter Four

The aim of this Chapter is to examine the construct validity of the DERS-R in relation to psychological distress, namely depression and anxiety using the K10 scale,

problem solving ability using the SPSI-R, and the personality trait of neuroticism using the IPIP. In order to provide evidence for construct validity the following relationships are expected:

- 1. The DERS-R will correlate with the K10 in a positive direction.
- Both the DERS-R and the K10 will correlate with 'negative orientation' on SPSI-R in a positive direction, and in a negative direction on 'positive orientation'.
- The DERS-R will correlate with 'neuroticism' in a positive direction, as measured by the IPIP.
- 4. The DERS-R will correlate with 'extroversion' in a negative direction, as measured by the IPIP.

4.4 Methods

A detailed outline of Methods for each sample and the instruments administered are reported in Chapter 2. However a brief review of the two samples and the questionnaires relevant to this Chapter are outlined below.

Sample 1 (n=327) comprised of community volunteers sourced from the general community, and the university. Individuals with a diagnosis of schizophrenia were accessed from the Schizophrenia Research Registry (n=129). These groups will be combined and referred to a combined sample.

Sample 2 comprised of first year university psychology students (n=264) who were offered course credit for taking part in the study.

4.4.1 Measures

In additional to the DERS-R three measures were used and will be summarised here (refer to Chapter 2 for extensive details of each measure):

- 1. K10 is a measure of psychological distress (Kessler et al., 2002).
- 2. SPSI-R is a measure of social-problem solving ability with a focus on two aspects; problem-solving skills and problem orientation (Thomas J. D'Zurilla et al., 1998).
- 3. IPIP is a five factor model of personality traits. (Goldberg et al., 2006).

Sample 1 was administered the DERS-R, K10 and SPSI-R and sample 2 was administered the DERS-R, K10 and IPIP.

4.4.2 Statistical analysis

Descriptive statistics are reported on both samples. In order to assess construct validity, correlations were conducted to determine the association of the DERS-R with other measures.

4.5 Results

4.5.1 Construct Validity of the DERS-R

4.5.1.1 Descriptive Statistics

The clinical and non-clinical sample (Sample 1) consisted of 260 participants comprising 79 males. The age range was as follows: category 18-40 years, n=161; and greater than 40 years, n=93 (six participants failed to record age). The total DERS-R mean=74.92, SD=23.16. The total mean reported on the K10=17.94, SD=6.83.

Sample 2 consisted of 37 males, and 186 female university students. The age range reported by n=223 was between 18-40 years, mean=22.60, SD=7.70. The total DERS-R mean=85.14, SD=22.57. A total K10 mean=22.16, SD=7.75 was reported.

To determine whether the difference in total DERS-R mean scores between the two samples was significant a t-test was conducted. This analysis revealed a statistically significant difference between the samples, with the university students scoring significantly higher on the DERS-R compared to the combined sample, $t_{(236)}$ =6.95, p < .00. There was also a significant difference between the two samples in the K10 scores, $t_{(236)}$ =8.84, p < .00, indicating the university students experienced greater psychological distress.

4.5.1.2 DERS-R and K10 relationship

To determine if the DERS-R correlated with the K10 in a positive direction, correlations were conducted between the two measures using samples 1 and 2. Table 4-1 shows a positive correlation in the expected direction between the DERS-R and K10 scores, so that an increase in scores on the DERS-R was associated with an increase in scores on the K10, in other words as individuals experienced greater difficulties in ER they also experienced greater psychological distress. When analysis was repeated in

sample 2 (see Table 4-2), a significant positive correlation between the DERS-R and K10 scores was also revealed in the expected direction.

Sample 1	DERS-R	K10	Factor 1 Strategies	Factor 2 Awareness	Factor 3 Non-Acceptance
K10	.741**				
Factor 1 (Strategies)	.785**	.676**			
Factor 2 (Awareness)	.536**	.297**	.232**		
Factor 3 (Non- acceptance)	.743**	.468**	.384**	.274**	
Factor 4 (Goals)	.713**	.548**	.469**	.203**	.396**

Table 4-1 Correlations between scales and factors in combined clinical/non-clinical sample

** Correlation is significant at the 0.01 level (2-tailed).

* Correlation is significant at the 0.05 level (2-tailed).

Table 4-2 Correlations between scales and factors in student sample

Sample 2	DERS-R	K10	Factor 1 Strategies	Factor 2 Awareness	Factor 3 Non-Acceptance
K10	.611**				
Factor 1 (Strategies)	.786**	.489**			
Factor 2 (Awareness)	.452**	.192**	.102		
Factor 3 (Non- acceptance)	.756**	.449**	.456**	.201**	
Factor4 (Goals)	.650**	.430**	.462**	.061	.312**

** Correlation is significant at the 0.01 level (2-tailed).

* Correlation is significant at the 0.05 level (2-tailed).

4.5.1.3 Component relationship of Factor Analysis Item Loadings

In Chapter 3 the final factor analysis of the DERS-R indicated four factors or components to be the optimal. These were referred to as 'strategies', 'awareness', 'nonacceptance', and 'goals'. The item loadings in the factor analysis on component three (non-acceptance of emotion responses) were all negative, hence, for the purpose of interpretation, we reversed the signs so that high scores on factor three indicated acceptance.

Table 4-1 and Table 4-2 show the correlations for the four factors retained in the DERS-R. The combined sample in Table 4-1 shows statistically significant moderate to high correlations on all four factors of the DERS-R in the expected direction. All four factors also had a statistically significant low to moderate correlation with each other in the expected direction.

The DERS-R component one 'strategies', items reflect a limited access to strategies and showed a significantly positive relationship with 'awareness', which reflects a lack of 'awareness' to emotion responses. 'Strategies' were also significantly, and positively related to 'goals' where items reflect a limited access to 'strategies' making it difficult to concentrate and accomplish tasks when experiencing negative emotions. The last component labelled 'non-acceptance', showed a significantly positive relationship to 'strategies', indicating if someone has limited access to strategies when upset then they would find it difficult to accept their initial negative reactions.

Table 4-2 represented the student sample and showed a similar result when compared to the combined sample. As expected, the DERS-R scores produced statistically significant moderate to high correlations on all four components in the expected direction. Component one showed a statistically positive relationship with 'nonacceptance' (r=.46), and 'goals' (r=.46) and a non-significant relationship with 'awareness' (r=.10). Awareness was also not-significantly correlated with 'goals' (r=.06).

4.5.1.4 Sample 1: DERS-R and SPSI-R

In order to determine the relationship between the DERS-R total score and the SPSI-R (which has two dimensions, positive orientation and negative orientation), correlations were conducted using Sample 1. Analysis revealed a significantly moderate negative relationship between the DERS-R score and positive orientation (see Table 4-3). That is, low scores on positive orientation, reflecting poor problem solving ability, were associated with a high total score on DERS-R indicating difficulties in ER. Thus, poor problem solving was associated with higher difficulties in ER.

A significantly positive, strong relationship was found between the SPSI-R negative orientation and the DERS-R. This indicated that those experiencing greater difficulties in ER, also reported high scores on SPSI-R, suggestive of an inhibitive cognitive-emotional style.

As expected, there was a weak negative association between positive orientation and negative orientation (see Table 4-3), indicating participants who received low scores on positive orientation reported higher scores on negative orientation. In other words participants who scored low on positive orientation, indicating poor problem solving ability, tended to score highly on negative orientation, indicating inhibitive cognitive-emotional style.

Sample 1	DERS-R total	K10	Positives SPSI-R	Negatives SPSI-R
Positives	344**	308**		
Negatives	.744**	.643**	300**	
Total SPSI	732**	635**	.644**	913**

Table 4-3 Correlations between the DERS-R and SPSI-R in a combined clinical/non-clinical sample.

** Correlation is significant at the 0.01 level (2-tailed).

* Correlation is significant at the 0.05 level (2-tailed).

The SPSI-R total score was negatively related to the K10, indicating high scores on psychological distress were related to lower scores on SPSI-R total score. In other words, participants who were experiencing high distress also had difficulties in problem solving. The K10 significantly and positively correlated with the negative orientation on the SPSI-R, indicating high scores on psychological distress (K10) were associated with high scores on the SPSI-R, reflecting inhibitive cognitive-emotional style (see Table 4-3). The K10 significantly and positively correlated with the negative orientation on the SPSI-R, indicating high scores on psychological distress (K10) were associated with high scores on the SPSI-R, reflecting inhibitive cognitive-emotional style (see Table 4-3). The K10 significantly and positively correlated with the negative orientation on the SPSI-R, indicating high scores on psychological distress (K10) were associated with high scores on the SPSI-R, reflecting inhibitive cognitive-emotional style.

4.5.1.5 Sample 2: DERS-R and Big 5 Personality Traits

Correlations were used to determine the relationship between the DERS-R total score and the Big 5 Personality Traits scale using the student data (Sample 2). Analysis on the student sample revealed a significantly higher positive relationship on the neuroticism subscale (refer to Table 4-4), with high scores on DERS-R. In other words, the more one is emotionally dysregulated the higher the score on the neuroticism scale indicating that one experiences greater worry, nervousness and a sense of insecurity. A statistically significant, but weak, negative relationship was found for the subscales of extroversion, openness and agreeableness with a moderately negative association on the conscientiousness subscale (see Table 4-4).

	DERS-R	Neurotic	Extroversion	Open	Agree	Consciousness
Neurotic	.680**					
Extroversion	286**	378**				
Open	136*	078	.311**			
Agree	209**	098	.098	.296**		
Consciousnss	422**	394**	.116	.142*	.305**	
K10	.611**	.605**	258**	093	190**	283**

Table 4-4 Correlations between the DERS-R and Big 5 Personality Traits in sample 2

 ** Correlation is significant at the 0.01 level (2-tailed).

* Correlation is significant at the 0.05 level (2-tailed).

As expected, extroversion correlated with both the K10, r=-.26, p < 0.1, and the DERS-R, r=-.29, p < 0.1 in the expected direction. This indicates that the more one is psychologically distressed, and experiencing greater difficulties in ER, the less likely one is to score highly on the extroversion scale, indicating sociability, talkativeness and a fun-loving and affectionate nature.

4.6 Discussion

The purpose of this Chapter was to examine the construct validity of the DERS-R by investigating the relationship between the DERS-R and other measures of psychological distress as measured by the K10. Construct validity was supported by a significantly high correlation between the DERS-R and the K10 in the expected direction. This finding was evident in both samples, and is consistent with studies that report individuals who experience difficulties in ER are also more likely to experience high psychological distress (Andrews et al., 2000; Henderson, Andrews, & Hall, 2000).

It is interesting to find that the combined sample (Sample 1) were not only seen as statistically different from the student sample (Sample 2), they also reported experiencing less psychological distress and less difficulties in ER. It would be generally considered that a clinical sample with a diagnosis of schizophrenia would experience greater psychological distress compared to healthy controls. The rationale for combining the clinical (schizophrenia) sample with a non-clinical sample was outlined in Chapter 2. In summary, the combining of the samples was thought to be more representative of the general population when compared to a student population. One possible explanation for higher distress scores in Sample 2 has been presented by Stallman (2010) who conducted a study with over 6000 students from two Australian universities and reported 83.9% of students experienced elevated distress levels as measured by the K10. The proportion of students who reported in the severe range for anxiety and mood disorders was 19.2%. This result is consistent with results found in the general Australian population reported by Andrews et al (2002).

As expected the results showed both the DERS-R and the K10 were associated with 'negative orientation' of inhibitive cognitive-emotional style on the SPSI-R in a positive direction. This suggests that the more one is experiencing psychological distress and difficulties in ER, the greater the tendency to doubt one's ability to solve problems successfully and the tendency to become easily frustrated when confronted with problems in living. This is supported by two studies, one by D'Zurilla et al (2002), and another by Nezu et al (1985). Both studies found that psychological distress was correlated with ineffective problem solving.

The validity of the DERS-R was further supported, with a strong convergence between the DERS-R and 'neuroticism' on the IPIP scale in Sample 2. This indicates the more one is experiencing difficulties in ER the higher the scores on the 'neuroticism' scale. This finding is supported by Weiting et al (2009) who conducted two studies examining the relationship between personality differences in emotions. They found neuroticism was associated with increased negative emotions, and more specifically, that the difference in the emotions experienced can result from a difference in the tendencies to regulate emotions. Another study by Davis et al, (1998) also reported that a tendency to reduce or eliminate one's negative emotion correlated negatively with neuroticism.

On the basis of the current findings, it seems reasonable to conclude that the DERS-R has good construct validity and is a suitable scale to use in clinical samples. Further research is needed to examine whether the DERS-R can show change after a clinical intervention. It is possible for instruments to be reliable, but unresponsive to change, so in order for the DERS-R to be an effective instrument for use in a clinical setting it needs to be able to show change in a population after they have received a clinical intervention. Chapter 5 will examine the sensitivity of the DERS-R using a series of repeated measure ANOVAs. It will compare change in a clinical sample after receiving a treatment intervention, at baseline and 3, 6 and 12 months follow-up.

94

CHAPTER 5 THE DERS-R SENSITIVITY TO CHANGE

5.1 Overview

It is common knowledge that a standardised measure can report sound reliability and validity, yet be regarded as ineffective in a clinical setting, unless it is able to detect change pre- and post- an intervention. In other words, a standardised measure has greater feasibility or validity if it has been tested on a clinical population and has proven to be sensitive to change post treatment.

This Chapter will examine the DERS-R sensitivity to change in a clinical population using the same clinical group reported in Chapter 4; CDSA (co-existing diagnosis of depression and substance abuse). As not all participants from the CDSA data, scored above threshold on all substances (NHMRC, 2004) only those who scored above threshold for alcohol use, and those who scored above 17 on the Beck Depression Inventory were included (1996), (Pols & Hawks, 1992). Participants were then randomly allocated to one of three interventions. Measures were taken at baseline, and at 3, 6, and 12 months following baseline. All active treatments had a similar number of sessions at the design level, and there was no control group in this analysis. The study did not examine treatment effects, but, did examine the effects of treatment allocation with results indicating treatment allocation did not impact on change scores in depression or alcohol use. Change was determined by a significant change, in either direction, in alcohol consumption or depression scores pre – post- intervention. The DERS-R also detected change on both variables.

5.2 Introduction

This research has shown that the DERS-Revised (DERS-R, 29 items) has sound test-retest reliability, and can be considered as an assessment instrument with stability over time. The next step would be to determine whether the DERS-R is sensitive enough to detect a change in participants' responses as a result of receiving an intervention. Sensitivity to change can be defined as the ability of a measure to detect clinically relevant changes over time (Guyatt, Walter, & Norman, 1987).

It could be suggested that an outcome measure's sensitivity to change is directly related to the ability of the instrument to do what it purports to do, which is to measure an individual's change in symptoms and behaviour over time due to an intervention. Therefore, the concept of sensitivity to change can be conceptualized as an issue of construct validity (Vermeersch et al., 2000). Two aspects of construct validity are referred to as convergent and discriminant validity. Convergent validity is where the scores between two theoretically related measures are significantly correlated such as between the DERS-R and BDI-II. Discriminant validity refers to a measure's ability to discriminate between constructs that are theoretically different. Predictive validity can be referred to as construct-related discriminant validity (Nunnally et al., 1994; Gandek, WareJr et al., 1998). More specifically, predictive validity is used to determine the degree to which scores predict outcome on a future criterion. For example, the predictive validity of health status measures can be tested by predicting some future event such as job loss or utilization of health care services (Nunnally et al., 1994). Similarly, a measure of difficulties in ER that can track the change of participant's progress, in a clinical population pre-post intervention, may be associated with the efficacy of treatment.

A number of criteria have been suggested when looking for sensitivity to change in an instrument. First, the scores should reflect change as a result of an intervention and this change must occur in the theoretically proposed direction. Secondly, scores should not be attributable to measurement error (Vermeersch et al., 2000).

Jacobson and Truax (1991) have outlined three criteria that might be considered when examining clinical significance:

- 1. The level of functioning subsequent to therapy should fall outside the range of the dysfunctional population, where range is defined as extending to two standard deviations beyond (in the direction of functionality) the mean for that population.
- 2. The level of functioning subsequent to therapy should fall within the range of the functional or normal population, where range is defined as within two standard deviations of the mean of that population.

3. The level of functioning subsequent to therapy places that client closer to the mean of the functional population than it does to the mean of the dysfunctional population (p13).

This has been supported by other researchers who have suggested that for a measure to be useful, the items should be sensitive to symptomatic and behavioural changes that occur with treatment (Lambert & Hill, 1994, Vermillion #1057). This statement seems similar to criterion three outlined by Jacobson and Truax's (1991). However, it is worth mentioning that symptomatic and behavioural change can increase or decrease as result of treatment due to unexpected responses or reactions a client has to the treatment. Therefore, change might be indicated by scores moving away from the mean of the functional population. When determining whether an instrument can detect or is sensitive to change we can measure change in either direction, either closer to the mean or further away from the mean. This Chapter will examine if there is a change in alcohol consumption and depression scores as a result of an intervention, and whether the DERS-R can detect this change.

In Chapter 1 it was reported that individuals who experience depressive episodes tend to have difficulties in ER (Kring & Werner, 2004; Campbell-Sills et al., 2006; Garnefski et al., 2006). However, there have only been a few studies that have reported a change in both mood and DERS (36-item) scores post intervention. Using the DERS as a measure, one study by Ehring (2008) recruited university students, half of whom reported they had experienced a past major episode of depression. The study excluded any individual with a score higher than 11, indicating no participant was 'currently depressed' as scored by the BDI-II (Beck et al., 1996). A comparison of the never depressed and recovered depressed groups found evidence to suggest that depression vulnerability was related to deficits in ER as measured by the DERS (Gratz 2004 #225).

Recruiting a clinical population with a diagnosis of borderline personality disorder (BPD), another study reported a significant and progressive improvement in both mood (measured by Depression and Anxiety Scale (DASS) (Lovibond & Lovibond, 1995) and emotion dysregulation (measured by the DERS) following a 3 month, short-term treatment, which comprised of partial hospital and intensive

97

outpatient levels of care (Gratz, Lacroce et al., 2006). While the above study is promising the main limitation was the small sample size at baseline to one month (n=29), and baseline to three months (n=18) which makes it difficult to generalise the results. Similar results were reported in a study that utilised a 14 week group intervention designed to teach women with a diagnosis of BPD more adaptive ways of responding to their emotions. This study found a significant reduction in both depression (measured by DASS), and emotion dysregulation (measured by the DERS). This study, however, was also limited by a small sample size (group and treatment as usual n=12, and treatment as usual n=10) (Gratz & Gunderson, 2006).

To determine if the DERS-R is sensitive to change post-intervention, this research will employ a clinical population, all of whom score above threshold (17 or greater) on the BDI-II depression inventory, indicating moderate to severe depression (Beck et al., 1996), and all scored over threshold for alcohol abuse.

5.2.1 Alcohol, depression and emotion dysregulation

The Diagnostic and Statistical Manual of the Mental Disorders [DSM-IV] (A.P.A., 2000) shows 75% of the categories of psychopathology are characterised with emotion or ER problems. Emotion dysregulation, especially in children, can be so prominent that the disorders are defined primarily on the basis of disturbed emotions (Mineka & Sutton, 1992).

Of the mental disorders that are related to alcohol dependence, up to 80% complain of having a depressive episode or major depression (Regier et al., 1990; Kessler et al., 1996; Kessler et al., 1997). Research documenting the comorbidity of alcohol and depression is prevalent (Khantzian, 1990; Regier et al., 1990; Penick et al., 1994; Strakowski, Tohen, Flaum, & Amador, 1994; Raimo et al., 1998; Teesson et al., 2000; Patrick, 2003; Sullivan et al., 2005; Burns & Teesson, 2006).

The same clinical group (CDSA) that was used in Chapter 4 of this research will be employed for this study. However, not all participants used all substances. Therefore only participants who reported alcohol use above threshold were selected.

This Chapter will determine if the DERS-R can detect change after an intervention as per Jacobson and Truaz (1991). It will determine whether there is a change in the clinical groups as a result of treatment, by looking for a change in the

mean scores in the DERS-R, the BDI-II, and the consumption of alcohol, in the expected direction from baseline and over the 3 time periods. Second, if there is a change in mean scores, it will determine if this can be detected by the DERS-R. Change as a result of treatment will be determined if there is a reduction in BDI-II (depression scores), and alcohol use that will be scored by the Drug Use Scale of the Opiate Treatment Index (OTI) (Darke et al., 1991). DERS-R sensitivity to change will be reflected by a change in scores in both QTI scores for alcohol and in BDI-II. It is expected that participants will score closer to the mean of a functional population in both alcohol consumption and depression scores after receiving an intervention compared to the mean at baseline.

5.3 Aim of Chapter Five

The aim of this Chapter is to examine the ability of the DERS-R to determine if it can detect change following a treatment intervention in a clinical population.

5.4 Methods

5.4.1 Participants and procedure

A clinical group with diagnosed co-existing depression and alcohol abuse, (n=103), was invited to be randomly allocated to one of three treatment trials: therapist, computer or brief intervention. The treatment study administered the DERS (Gratz et al., 2004) on four assessment occasions: baseline, 3, 6, and 12 month follow-up. The following measures were also used as part of this study to assist in determining whether there was a change as a result of treatment and if this change was reflected in the DERS-R. A detailed outline of these measures has been reported in Chapter 2, however a summary is provided below.

5.4.2 Measures

1. The Drug Use Scale of the Opiate Treatment Index (OTI) (Darke et al., 1991)

The Drug Use Scale of the OTI reports on the quantity and frequency of substance use across 11 drug types including: alcohol, cannabis, heroin, other opiates, amphetamines, cocaine, hallucinogens, barbiturates, tranquilisers, inhalants and tobacco. Each drug type is individually assessed, and clients report on the last three occasions of use in the month prior to assessment, estimating the number of units consumed on each of these occasions. The scales on the OTI can be used as a whole, or individually without compromising the validity or reliability of the scale. In this study only the alcohol score is reported as it was the predominant substance use problem.

2. Structured Clinical Interview for DSM-IV (SCID) (First et al., 2001)

The SCID is a semi-structured interview and provides a diagnostic clinicianrated measure of alcohol abuse and dependence based on the criteria set in DSM-IV (A.P.A., 2000). It relies in part on clinician judgement to derive a diagnosis, therefore, reliability of the scale is related to the context in which it is being used (First et al., 2001).

3. Beck Depression Inventory-II (BDI-II) (1996)

The BDI-II is commonly used to screen for depressive symptoms among people with drug and alcohol use problems (Dawe et al., 2002). It is a 21-item self-report questionnaire used to screen for depressive symptoms over the previous two-week period, and takes five minutes to complete. The BDI-II has good internal consistency among psychiatric outpatients (a=0.93) and with a non-clinical sample (α =0.93).

4. Clinical Intervention

The treatment interventions were originally described in a study by Kay-Lambkin (2011). A brief summary is reported here. All participants received one faceto-face treatment session and were then randomised to receive either: an additional (i) nine sessions of integrated cognitive behaviour therapy and motivational interviewing (CBT/MI) delivered by a therapist; (ii) nine sessions of integrated (CBT/MI) delivered by a computer, with minimal therapist assistance (Computer Assisted Counselling or CAC); or (iii) nine sessions of supportive counselling delivered by a therapist (personcentred therapy, PCT). Randomisation was concealed from the therapist and participants until the conclusion of session 1. Studies using a similar sample have shown both CBT and MI treatments to have efficacy in reducing alcohol use. For a comprehensive review refer to the following studies: (Baker, Turner, Kay-Lambkin, & Lewin, 2009; Kay-Lambkin et al., 2009; Baker et al., 2010; Kay-Lambkin, Baker, Lewin et al., 2011; Kay-Lambkin, Baker, Kelly et al., 2011).

Each intervention was manualised. The nine sessions of Cognitive Behavioral Therapy/Motivational Interviewing (CBT/MI) involved 9 x 60 minute sessions of CBT,

with MI employed thematically throughout treatment; and the content for CAC was identical to the content in the therapist-delivered integrated CBT/MI, however it was delivered via computer. Therapist input was a brief, structured 10 minute check-in session at the conclusion of each session. The PCT was adapted from Sellman, Sullivan, and Dore (unpublished): Brief Treatment Programme for Alcohol Dependence; Person-Centered Therapist Manual, Christchurch School of Medicine, New Zealand. It was included to control for therapist (live) contact, but not for therapy content.

All active treatments had a similar number of sessions at the design level, and there was no control group. The major analysis in this study focussed solely on whether the DER-R detected change between baseline and follow-up. This treatment program was designed to reduce alcohol consumption as oppose to specifically targeting difficulties in emotional regulation. It will be interesting to determine if there is a change in alcohol and depression levels, whether the DERS-R is able to detect this change when a non-specific treatment program is conducted. It also did not examine whether one treatment was more effective in behavioural change over another. However it did examine whether treatment allocation was associated with change.

5.4.3 Statistical analysis

Data were analysed using the Statistical Package for Social Sciences (PAWS 18) for Windows. Mean scores and standard deviations (SD) are reported for the DERS-R, BDI-II and alcohol scores. To establish whether differences in the mean scores on each measure were significant a series of one-way repeated measures analysis of variance (ANOVA) were conducted. Change in mean scores was analysed by ANOVA with treatment groups in order to determine whether change was a result of a specific treatment allocation. In all repeated measures ANOVA, the multivariate test statistic was used, as it does not require the assumption of sphericity.

Trend analysis will be reported to determine direction of the DERS-R mean scores. In order to further reduce error, a random effects model was used to analyse measures over four time periods.
5.5 Results

5.5.1 DERS-R, BDI-II and OTI alcohol mean scores over time of treatment

DERS-R mean scores for each time period are presented in Table 5-1 and indicate a reduction in the mean scores over time. In order to establish whether this reduction is significant, a one-way repeated measure ANOVA was conducted to compare the DERS-R mean scores at each time-period: baseline, 3, 6 and 12 months follow-up.

There was a significant effect in DERS-R total scores over time, indicating that the DERS-R total was significantly lower at each of the 3, 6, and 12 month follow-up assessments relative to baseline [F (1, 2) 1928.28, p<0.001. These results show a change in DERS-R mean scores over time regardless of treatment allocation [F (1, 2) 1.35, p< 0.26]. However there is no indication that this change is associated with the specific treatment that participants received.

BDI-II mean scores for each time period are also presented in Table 5-1 and indicate a reduction in the mean scores over time. A one-way repeated measure ANOVA was conducted to compare the BDI-II mean scores at each time-period; baseline, 3, 6, and 12 months follow-up. There was a significant effect in BDI-II total mean scores over time, indicating that the BDI-II total was significantly lower at each of the 3, 6, and 12month follow-up assessments relative to baseline, [F(1, 2) 691.58, p<0.001], and this was not a function of treatment allocation [F(1,2) 1.02, p < 0.47].

				Planned Trend Contrasts		
Measure	Ν	Time	Mean (SD)	Linear Component	Quadratic component	Cubic Component
DERS-R	86	Baseline	85.12 (16.81)	F(1, 85) 34.41,	F(1, 85) 9.36,	F (1, 85) 6.04,
	86	3 months	74.62 (19.95)	p <0.001	p < 0.003	p < 0.02
	86	6 months	74.66 (18.94)			
	86	12 months	72.70 (20.13)			
BDI-II	95	Baseline	31.86 (8.74)	F(1, 94) 97.25,	F(1, 94) 39.13, p < 0.001	F(1, 94) 3.47, p < 0.07
	95	3 months	21.66 (11.60)	p <0.001		
	95	6 months	19.93 (11.43)			
	95	12 months	20.98 (10.77)			
OTI (Alcohol)	103	Baseline	7.66 (7.35)	F (1, 102) 21.65,	F (1, 102) 14.54, p < 0.001	F (1, 102) 3.88, p < 0.052
	103	3 months	4.07 (5.40)	p < 0.001		
	103	6 months	3.99 (5.11)			
	103	12 months	4.22 (6.31)			

Table	e 5-1	Mean	scores.	SD.	orthogona	l pol	vnomial	contrasts
				- ,				

* Where subtotals < 103, missing data due to self administration

To determine whether there was a reduction in alcohol mean scores over the time of intervention, the mean scores for alcohol use at baseline, 3, 6, and 12 months were calculated. Table 5-1 indicates alcohol total mean scores had decreased from baseline to 3 months, and this reduction tended to remain with little change from 3, 6 and 12 months. A repeated measure of ANOVA was conducted to compare mean scores on alcohol use at baseline, 3, 6 and 12 months follow-up. There was a significant reduction in alcohol use, indicating that alcohol consumption was significantly lower at each of the 3, 6 and 12 month follow-up assessments relative to baseline [F (1,2) 105.12, p< 0.000] and this was not a function of treatment allocation [F (1,2) 0.509, p<0.60.

5.5.1.1 Association between depression and DERS-R scores

Correlations were conducted between the DERS-R baseline, 3, 6, and 12 month scores and the BDI-II depression scores at 3, 6, and 12 month. There was a positive correlation found between DERS-R and BDI-II scores, indicating that as scores on DERS-R decreased, scores on the BDI-II also decreased (Table 5-2).

	BDI-II	3 month-BDI-II	6-month BDI-II	12-month BDI-II
DERS-R baseline	.396**	.185*	.314**	.211*
DERS-R 3 month	.257**	.593**	.524**	.389**
DERS-R 6 month	.419**	.526**	.680**	.502**
DERS-R 12 month	.255**	.388**	.391**	.660**

Table 5-2 Correlations between DERS-R and BDI-II

**. Correlation is significant at the 0.01 level (2-tailed).

In order to determine if a change in DERS-R scores corresponded with a change in depression scores, a new variable was created by subtracting 3 month BDI-II from baseline BDI-II; the same was done with the DERS-R scores, (subtracting the 3 month DERS-R from DERS-R baseline) such that positive scores in the new variable was indicative of improvement over that time-period. New variables were created for 6 and 12 month scores in both the BDI-II and DERS-R in the same manner.

	BDI-II baseline - 3 months	BDI-II baseline - 6 months	BDI-II baseline - 12 months
DERS-R baseline and 3 months	.466**	.384**	.319**
DERS-R baseline and 6 months	.379**	.484**	.312**
DERS-R baseline and 12 months	.302**	.246**	.565**

** Correlation is significant at the 0.01 level (2-tailed).

Table 5-3 shows positive correlations at a significant level such that an improvement in difficulties in ER correlated with an improvement in depression.

5.5.2 Trend analysis

Uncorrelated polynomials are a more complex design and mean that the linear, quadratic, cubic components are partitioning distinct, or different parts, of the variation in the data. First degree polynominals are indicated by a constant decline or significant linear trend. If the change is slow over the first few timeframes and drops sharply over the remaining time-periods a quadratic trend, or second degree polynomial will be found. Finally if the decline is slow at the first timeframe then drops sharply and finally levels off, a cubic trend or third degree polynomial will be found (Stevens, 1996, p460) A trend analysis indicated the difference between the DERS-R mean scores over each time period were significant on linear, quadratic and cubic components with the linear trend most pronounced [F(1,85) 34.41, p <0.001]. The difference between the BDI-II mean scores over each time period, was significant for both linear and quadratic components at .001 level. The difference between the QTI alcohol mean scores over each time period were significant at linear, and quadratic components at 0.001 level, and cubic at 0.052 level (refer to Table 5-1).

5.5.3 Random effects model

To reduce standard error, a random effects model was undertaken to examine the association between BDI-II and DERS-R at each time period of baseline, and 3, 6, 12 month follow-up. This showed a highly statistically significant association (p<0.00) and a coefficient of .89, indicating that each one unit increase in the BDI-II is associated with an 0.89 unit increase in the DERS-R (note: as mentioned earlier, positive scores in the new variable were indicative of improvement). The within person R-Squared is 0.38 and the between person R-squared is 0.42 indicating there is a strong association of BDI-II and DERS-R both within person and between person.

Finally, a random effects model was also undertaken to examine the association between DERS-R and alcohol at each time period. This showed a highly significant association (<0.001) and the coefficient of 0.64 indicates that each one unit increase in alcohol is associated with an increase of 0.64 unit in the DERS-R. The within person Rsquared is .08 and the between person R-squared is .09 indicating a moderate association of DERS-R and alcohol use both within person and between person.

5.6 Discussion

This Chapter set out to examine the DERS-R sensitivity to detect change prepost a clinical intervention, in a clinical population. The first aim was to determine if both the DERS-R and BDI-II showed a change in mean scores at baseline, and 3, 6, 12 months post treatment intervention, and if this change was significant. The DERS-R was shown to be sensitive to change after intervention: both when analysing the mean scores and when correlating baseline, and 3, 6, and 12-month follow-up changes on this scale with those changes in another measure, namely BDI-II. This is interesting, as the intervention did not specifically aim to address difficulties in ER. Rather, the intervention was aimed at reducing depression and alcohol consumption. It could be argued that CBT has a focus on the unhelpful beliefs about emotions, which contribute to the role of distress and impairment. However, not all interventions used in this study contained CBT as an intervention.

As expected, the DERS-R and BDI-II were found to be highly correlated at each time period. In other words, as depression improved participants experienced fewer difficulties in ER. While it was not the intention to show convergent validity of the DERS-R in this Chapter, the significant correlations between these two instruments add further support to the suggestion that they are theoretically related constructs.

The third aim was to determine if there was a change in alcohol consumption over time and a corresponding change in the DER-R. Mean scores indicated a reduction in the mean scores on alcohol use, with a noticeable move closer to the mean of a functional population. Most importantly, the DERS-R was able to detect this change, making it a useful measure which is sensitive to symptomatic and behavioural changes that occur with treatment (Jacobson et al., 1991; Vermillion & Pfeiffer, 1993; Lambert et al., 2004).

Several limitations warrant mention. First, the effectiveness of the treatment interventions was not assessed as it did not have a direct bearing on the goal for this Chapter. While change has occurred and the DERS-R is able to detect this change, it has not been determined if a specific arm of the intervention in the RCT had greater success in reducing emotion dysregulation. Furthermore, the absence of a control group prevent controlling for the effects of time on symptom change. There is evidence to suggest that PCT used as a supportive therapy is a therapy in itself, similar to a number of psychodynamic, or self psychotherapies that do not directly target the problematic behaviour (Tobin, 1991; Portnoy, 1999). Therefore PCT cannot be used as a control group for the purpose of this Chapter. Using PCT as an intervention is not the same as a control group where participants have baseline assessments and are placed on a wait list for 3 months prior to starting treatment. As a result, sensitivity to change results have been explored in a more general sense, without claim that psychotherapy was the mechanism of the observed changes.

Second, this study only used one additional measure to assess emotion change, that is, the BDI-II, and while it proved to be a good measure of convergent validity, it would be interesting to see if other depression measures reveal similar results. Finally, participants recruited had a personal interest in accessing the free intervention to address their depression and alcohol consumption. This may have biased the result on participant self-reporting and on level of motivation to improve their situation. In other words participants who volunteer for an intervention may be more invested in improving their situation as opposed to participants who are non-volunteers.

Two major strengths of the DERS-R are that it is slightly briefer than the original DERS (Gratz et al., 2004) while retaining strong components. The DERS-R can also detect change in a clinical population with co-existing depression and alcohol use. These two issues are important to clinicians who are generally constrained for time and are constantly looking for a brief self-reporting measure that can be utilised in a clinical setting. Lambert et al (2004) indicates a measure that is complex, has awkwardly worded items, or is lengthy to complete, might lead to a client's refusal to complete. It can also be time consuming for the clinician to score and interpret leading to the abandonment of the measure, and therefore it has little clinical value.

Another major advantage of the DERS-R for clinicians is that many measures are designed for diagnosis as opposed to clinical effects of treatment on patients. A measure that shows sensitivity to change in clinical populations is considered one of the most important characteristics of any measure (Vermeersch et al., 2000; Hatfield et al., 2004; Lambert et al., 2004). There are a number of reasons reported as to why clinicians use outcome measures as part of their practice that vary from being able to track a client's progress; determine if there is a need to alter treatment; to maintain an ethical practice; and to determine the strengths and weaknesses of an intervention (Hatfield et al., 2004).

In conclusion, the quality of clinical outcome measures needs to be evaluated in a diagnosed clinical population for efficacy. This study found evidence of convergent validity. The DERS-R showed capacity to detect change in a clinical population post intervention, which is evidence of predictive validity.

CHAPTER 6 DISCUSSION

6.1 The DERS

ER is recognised as having a vital role in the development and maintenance of good mental health. When a person struggles to regulate their emotions they may report feeling 'out of control' and say or do things that they would not normally do. If this struggle to regain control, or regulate emotions continues, it can lead to experiencing difficulties in everyday life. An ongoing failure to regulate is then referred to as emotional dysregulation (ED). ED has been included in over half of the DSM-IV Axis I disorders and in all of the Axis II disorders (Gross & Levenson, 1997; Briere et al., 1998; Bushman et al., 2001; Gratz, 2003). While there have been numerous measures designed to capture various aspects of the ED construct, there are few that measure ED in adults.

One measure designed to capture ED was the Difficulties in Emotion Regulation Scale (DERS) (Gratz et al., 2004). Gratz and Roemer initially chose items to reflect difficulties within four domains: a) awareness and understanding of emotions; b) acceptance of emotions; c) the ability to engage in goal-directed behaviour and refrain from impulsive behaviour when experiencing negative emotions; and d) access to ER strategies perceived as effective. It was hypothesised that four factors would emerge from a factor analysis of the DERS. However, they found a six factor loading to be optimal. The six factors were referred to as 'non-acceptance', 'goals', 'impulse', 'awareness', 'strategies' and 'clarity'. The items thought to reflect 'awareness and understanding', loaded onto two separate factors and were named 'awareness' and 'clarity'. Similarly, items hypothesised to reflect goal directed behaviour loaded onto two separate factors and were named 'goals' and 'impulse'. Gratz and Roemer reported good reliability and validity of the DERS. The purpose of this thesis was to reexamine the psychometric properties of the DERS using data from an Australian community and clinically diagnosed participants.

6.1.1 Summary of the factor structure of DERS in this thesis

This thesis addressed two of the major limitations of the original study. First, as opposed to collecting data from university students, data for this thesis were collected from participants in the community and participants with a clinically diagnosed mental illness. Secondly, the size of the sample was increased from n=21 in the original study for the test-retest to n=557. Similar to the studies mentioned in Chapter 1 that indicated problems with the DERS subscales particularly the 'awareness' scale (Tull & Roemer, 2007; Weinberg et al., 2009), results from this study also indicated limitations of the six subscales. Four subscales or components were found to be optimal and reflected components that were similar to the hypothesised domains presented by Gratz and Roemer (2004); 'awareness', 'acceptance', 'goal', and 'strategies'. This thesis made several modifications to improve the model fit. Seven items were removed from the analysis and a component analysis was recomputed followed by a reliability analysis. The final four factor structure did not compromise the reliability of the scale. When examining the construct labels of the revised DERS (DERS-R), four of the original labels were retained: 'strategies', 'awareness', 'nonacceptance' and 'goals'.

6.1.2 Summary of construct, convergent, and predictive validity of the DERS-R

Chapter 4 explored the construct validity on the DERS-R, using the Kessler-10 (K10) (Kessler et al., 2003), the traits of neuroticism and extroversion on the Big 5 International Personality Item Pool (IPIP) (Goldberg, 1992), and the Social Problem Solving Inventory-R (SPSI-R) (1998) to examine predictive validity. Data from two samples were used. The first sample included community participants and participants who were clinically diagnosed with schizophrenia. The second samples were university students (a rationale for the two sample groups was presented in Chapter 4). T-tests revealed a statistically significant difference between the two samples on all measures. Significantly significant correlations on all the four factors of the DERS-R were found in both samples.

Construct validity was supported by a significant and high correlation between the DERS-R and the K10 in the expected direction, and in both samples. The relationship between DERS-R and SPSI-R in Sample 1 revealed a significant moderate, negative association between DERS-R scores and the 'positive orientation' subscale on the SPSI-R, and a significant strong, positive relationship between the DERS-R and 'negative orientation' on SPSI-R. This finding was similar in Sample 2. Correlations between the DERS-R and the IPIP in Sample 2 revealed a significant and strong positive relationship on the 'neuroticism' subscale and a statistically significant but weak negative relationship on the 'extroversion' subscale. These findings further supported the sound construct validity of the DERS-R.

It is known that an instrument can have sound validity and reliability yet be ineffective in a clinical setting. Therefore, Chapter 5 examined the DERS-R to determine whether the instrument was sensitive to change post treatment. Data were collected from a clinical population with clinically diagnosed co-existing depression and alcohol abuse, a subsample of the coexisting depression and substance abuse (CDSA) data. This subsample was randomised into one of three treatments, and the DERS-R was administered over four time periods: baseline, 3, 6, and 12 months follow-up. Analysis in Chapter 5 indicated the DERS-R was sensitive to change after a clinical intervention in a clinical population. In other words, the DERS-R was able to detect change, indicating that it is a useful measure, which is sensitive to changes that occur with treatment.

6.1.3 Limitations and future directions

There are several limitations to this thesis that need to be taken into consideration. While all clinical groups were formally interviewed, administered the SCID, and clinically diagnosed, all other measures were self-administered. Therefore the possibility of response biases, such as self-presentation or self-deception may jointly influence the measures used (Gratz & Gunderson, 2006). For example, the CDSA clinical group were volunteers who may have biased their responses in order to gain access to the treatment program being offered. While the researcher attempted to reduce this bias by randomising all participants into one of three programs, it is possible that participants, who are motivated to be relieved of their symptoms, might exacerbate their symptom reporting with the view that it would increase their chances of access.

In addition, these studies may be limited by the fact that the majority of participants in this thesis were recruited from the same city. However, participants were from four different sub-populations (combined community and university students, schizophrenia registry group, CDSA and a group containing only university students) allowing for larger numbers and more diverse populations than previously reported in other studies (Gratz et al., 2004; Gratz, Rosenthal et al., 2006; Fox et al., 2007; Fox et al., 2008; Gratz et al., 2008). Future research might address this limitation by including the DERS-R and conducting a multi-site study that includes a number of countries.

This thesis focussed on a few measures that were known to be associated with difficulties in ER, the K10 (anxiety and depression) BDI-II (depression), the SPSI-R (social problems), and the traits of neuroticism and extroversion on the Big 5 IPIP. Future research needs to consider administering other questionnaires of ER against which to evaluate convergent validity of the DERS-R. Furthermore, it is not known whether the symptomatologies precede and contribute to the development of difficulties in ER or whether experiencing difficulties in ER bring about other distressing symptoms. This thesis did not examine this issue. However, it is a topic of interest. Future research might include an experimental design, for example, using a clinically diagnosed depressed group and non-depressed group and examine ER longitudinally. Similarly a depressed group and a non-depressed group could be employed to explore difficulties in ER over time.

This research did not access participants who were clinically diagnosed with borderline personality disorders, and the DERS was initially developed among this population. However, this thesis aimed to extend existing research. Since the development of the DERS many studies have included the measure with populations other than borderline personality disorders (Salters-Pedneault et al., 2006; Fox et al., 2007; Tull & Roemer, 2007; Fox et al., 2008; Martin et al., 2011), which makes it timely that the psychometric properties of the DERS were re-examined among broader clinical groups. It is important that the measure functions well in various cultural settings and with various clinically diagnosed populations, as this has the benefit of allowing researchers and clinicians to assess similarities and differences in health impacts and to identify potential translation differences. It may also be said that one of the strengths of this thesis was that it included four different populations, two with a clinical diagnosis. The DERS-R continued to show sound psychometric properties. It is recommended that future studies should include the revised DERS-R when studying other clinically diagnosed populations, including borderline personality disorders.

While results in this thesis showed the DERS-R to be sensitive to change, the treatment that was administered in Chapter 5 was non-specific as it focussed on reducing depression and substance use behaviour, as opposed to targeting difficulties in ER. Future research utilising a treatment specifically designed to address difficulties in ER is needed to determine the DERS-R sensitivity to change in another setting. It would be interesting to conduct a randomised controlled trial in a diagnosed clinical population receiving treatment that will target ER specifically to determine if similar results are found.

6.1.4 Strengths of this research

The population size reported in each Chapter of this thesis met the required ratio of participants to items, and allowed for optimum statistical significance. This goes beyond many psychometric studies that report only partial aspects of their study, the majority of these being pilot studies as opposed to comprehensive psychometric studies on any measure. For instance Gratz (2004) accessed 357 university students, of whom 73% were female, to analyse the factor structure of the DERS and was found to have sound components. However, test-retest was conducted on a sample of only n=21.

Translational research can be referred to as an attempt to apply and test, in 'real world settings', clinical knowledge that has been acquired under research conditions (Schwartz, Trask, Shanmugham, & Townsend, 2004). There are a number of ways research can be referred to as translational as long as it meets the above requirement. In this thesis the use of a clinically diagnosed population to re-examine the reliability and validity of the DERS, along with an analysis of the measure's sensitivity to change using both a clinically diagnosed population and a treatment intervention population can be referred to as translational research.

The community population in this thesis was not restricted to university students, as it also included a general community sample thereby extending the demographics of the sample to include a wide and diverse age and gender group with various social and economic backgrounds. Collecting data from such a diverse population further strengthens the psychometric properties of the DERS-R as the validity of the factor structure and/or relationship between factors can alter when gender, culture, and other demographic characteristics are included in research samples. This research sought to minimise the impact of demographic variability by extending the sample size and characteristics of samples recruited.

Furthermore, by administering a non-specific treatment intervention in a clinical population, this research established the instrument's sensitivity to change. As a result it could be used to assess the efficacy of treatment outcomes on patients, regardless of treatment targets, making it a very useful instrument for clinicians as it could reduce the use of multiple measures.

Finally, the use of a number of different, clinically diagnosed populations to reassess the psychometric properties of the DERS-R demonstrates its ability to remain robust across populations with various diagnoses. Hence the DERS-R may be used in all clinically diagnosed patients to detect difficulties in regulation, and treatment efficacy regardless of the type of treatment intervention.

6.1.5 How the DERS-R can be used clinically

Standardised questionnaires can assist in clinical practice as they provide objective measures to evaluate clients' treatment progress. Such measures offer a degree of accountability, and quality assurance that subjective interviews fail to provide. With increased requirements for formal reporting on clients' progress by numerous funding bodies, court judges, solicitors, child protection, and claims to insurance and work-cover organisations, there is a need for reliable assessment that can measure change as a result of receiving treatment. Formal outcome assessment can assist the clinician in their judgement regarding focus of treatment, client progress, and client's functioning (Lambert et al., 1994; Lambert et al., 2004). One major reason cited by clinicians for not employing formal outcome measures is the time it takes to administer, score and enter the data (Hatfield et al., 2004). This suggests that a measure that is brief, taking less time for the clients to complete, and less time to enter and score the data, might be preferable to clinicians. Furthermore, a measure that has sound validity, reliability, and is sensitive to change post intervention is even more desirable to clinicians as it enables them to reliably track therapy progress, inform treatment decision making, determine if there is a need to alter therapy, and monitor the efficacy

of treatment (Hatfield et al., 2004). Outcome measures that are not psychometrically sound can provide information about symptom severity and client progress that is inaccurate. The ramifications of this are that clinicians might make decisions based on unreliable information, that can have detrimental outcomes for the client (Hatfield et al., 2004).

6.1.6 Summary

The DERS-R showed sound psychometric properties when used with clinical samples. The data supported the structure of the revised scale, which assessed four components of difficulties in ER: 'strategies', 'awareness', 'nonacceptance', and 'goals'. The results provided support for the validity of the measure, and also its sensitivity to detect change in participants who received a non-specific intervention.

REFERENCES

- A.P.A. (1994). *Global Assessment of Functioning Scale*. Washington DC: American Psychiatric Association.
- A.P.A. (2000). *Diagnostic and Statistical Manual of the Mental Disorders, (4th editionrevised)*. Washington DC: American Psychiatric Association.
- Aghevli, M. A., Blanchard, J. J., & Horan, W. P. (2003). The expression and experience of emotion in schizophrenia: a study of social interactions. *Psychiatry Res, 119, 261-270.*
- Allebeck, P. (1989). Schizophrenia: A life-threatening disease. *Schizophrenia Bulletin*, 15, 81-89.
- Anastasi, A. (1990). *Psychological Testing* (Sixth ed.). New York: Macmillan Publishing Company.
- Andreasen, N. C., & Olsen, S. (1982). Negative v positive schizophrenia. Definition and validation. *Archives General Psychiatry*, *39*, 789-794.
- Andrews, G., Sanderson, K., & Beard, J. (1998). Burden of disease. Methods of calculating disability from mental disorder. *British Journal of Psychiatry*, 173, 123-131.
- Andrews, G., Hall, W., Teeson, M., & Henderson, S. (1999). *The Mental Health of Australians*. Canberra: Mental Health Branch: Commonwealth Depertment of Health and Aged Care.
- Andrews, G., Hall, W., Teeson, M., & Henderson, S. (2000). Australia's mental health: an overview of the genreral population survey. *Australian and New Zealand Journal of Psychiatry*, 34, 97-205.
- Andrews, G., Henderson, S., & Hall, W. (2001). Prevalence, comorbidity, disability and sevice utilisation:Overview of the Australian National Mental Health Survey. *British Journal of Psychiatry*, *178*, 145-153.
- Andrews, G., & Slade, T. (2001). Intrepreting scores on the Kessler population. Australian and New Zealand Journal of Psychiatry Psychological Drestress Scale (K10). Australian and New Zealand Journal of Public Health, 25.
- Andrews, G., Slade, T., & Issakidis, C. (2002). Deconstructing current comorbidity: Data from the Australian National Survey of Mental Health and Well-Being. *British Journal of Psychiatry 18*, 306-314.
- Auquier, P., Lancon, C., Rouillon, F., Lader, M., & Holmes, C. (2006). Mortality in schizophrenia. *Pharmacoepidemiology & Drug Safety*, 15(), 873-879.
- Australian Bureau of Statistics. (2008). *The National Survey of Mental Health and Wellbeing (SMHWB)*. Retrieved. from.
- Ayuso-Gutierrez, J. L., & del Rio Vega, J. M. (1997). Factors influencing relapse in the long-term course of schizophrenia. *Schizophrenia Research*, *28*, 199-206.
- Babor, T. F., Higgins-Biddle, J. C., Saunders, J. B., & Monteiro, M. G. (2001). AUDIT. In *The Alcohol Use Disorders Identification Test: Guidelines for use in primary care.* Geneva: World Health Organisation, Department of Mental Health and Substance Dependence.
- Bagby, R., Parker, J. D., & Taylor, G. J. (1994). The twenty-item Toronto Alexithymia Scale: I. Item selection and cross-validation of the factor structure. *Journal of Psychosomatic Research*, 38, 23-32.

- Baker, A., Turner, A., Kay-Lambkin, F. J., & Lewin, T. J. (2009). The long and the short of treatments for alcohol or cannabis misuse among people with severe mental disorders. *Addictive Behaviors*, 852-858.
- Baker, A. L., Kavanagh, D. J., Kay-Lambkin, F. J., Hunt, S. A., Lewin, T. J., Carr, V. J., et al. (2010). Randomized controlled trial of cognitive behavioural therapy for coexisting depression and alcohol problems: Short-term outcome. *Addiction*, 87-99.
- Bardeen, J., R, Fergus, T., A, & Orcutt, H., K. (2012). An examination of the latent structure of the Difficulties in Emotion Regulation Scale. *Journal of Psychopathology and Behavioral Assessment*, 34, 382-392.
- Batel, P. (2000). Addiction and schizophrenia. *European Psychiatry*, 15, 115-122.
- Beck, A. T., Steer, R. A., & Brown, G. K. (1996). *The Beck Depression Inventory, Second Edition: Manual.* San Antonio: The Psychological Corporation.
- Beck, A. T., Steer, R. A., & Brown, G. K. (2000). *The Beck Depression Inventory Fast Screen Manual*. San Antonio: The Psychological Corporation.
- Berenbaum, H., & Oltmanns, T. F. (1992). Emotional experience and expression in schizophrenia and depression. *Journal of Abnormal Psychology*, 101, 37-44.
- Berking, M., & Wupperman, P. (2012). Emotion regulation and mental health: recent findings, current challenges, and future directions. *Current Opinions in Psychiatry*, 25, 128-134.
- Blanchard, J. J., Mueser, K. T., & Bellack, A. S. (1998). Andehonia, positive and negative affect, and social functioning in schizophrenia. *Schizophrenia Bulletin*, 24, 143-424.
- Bleuler, E. (1950). *Dementia Praecox of the group of schizophrenia* (J. Zinkin, Trans.). New York: International University Press.
- Bradley, S. (2000). *Affect Regulation and the Development of Psychopathology.* New York.: The Guildford Press.
- Brazelton, T., & Yogman, M. (1986). Affective development in infancy. New Jersey, USA: Ablex.
- Briere, J., & Gill, E. (1998). Self-mutilation in clinical and general population samples: Prevalence, correlates, and functions. *American Journal of Orthopsychiatry*, 68, 609-620.
- Briere, J. (2006). Dissociative symptoms and trauma exposure: Specificity, affect dysregulation, and posttraumatic stress. *Journal of Nervous & Mental Disease*, 194, 78-82.
- Brooks, R. T., Beard, J., & Steel, Z. (2006). Factor Structure and Interpretation of the K10. *Psychological Assessment*, *18*, 62-70.
- Bryman, A., & Cramer, D. (2001). *Quantitative Data Analysis with SPSS Release 10 for Windows*. London: Routledge.
- Burns, L., & Teesson, M. (2006). Alcohol use disorders comorbid with anxiety, depression and drug use disorders: Findings from the Australian National Survey of Mental Health and Well Being. *Drug and Alcohol Dependence*, 68, 299-307.
- Bushman, B., J, Baumeister, R. F., & Phillips, C., M. (2001). Do people aggress to improve their mood? Catharsis beliefs, affect regulation opportunity, and aggressive responding. *Journal of Personality and Social Psychology*, *81*, 17-32.

- Calkins, S. D. (1994). Origins and outcomes of individual differences in emotion regulation. *Monographs of the Society for Research in Child Development*, 59, 53-72.
- Campbell-Sills, L., Barlow, D. H., Brown, T. A., & Hofmann, S. G. (2006). Acceptability and suppression of negative emotion in anxiety and mood disorders. *Journal of Personality Assessment*, 54, 546-563.
- Campos, J. J., Frankel, C. B., & Camras, L. (2004). On the nature of emotion regulation. *Child Dev*, 75, 377-394.
- Carpenter, K. M., & Hasin, D. S. (1998). Reasons for drinking alcohol: Relationships with DSM-IV alcohol diagnoses and alcohol consumption in a community sample. *Psychology of Addictive Behaviors*, *12*, 168-184.
- Carpenter, W. T., Gold, J. M., Lahti, A. C., Queern, C. A., Conley, R., Bartko, J. J., et al. (2000). Decisional capacity for informed consent in schizophrenia research. *Archives of General Psychiatry*, 57, 533-538.
- Catanzaro, S. J., & Mearns, J. (1990). Measuring generalized expectancies for negative mood regulation: Initial scale development and implications. *Journal of Personality Assessment*, 54, 546-563.
- Catanzaro, S. J. (1997). Mood regulation expectancies, affect intensity, dispositional coping, and depressive symptoms: A conceptual analysis and empirical reanalysis. *Personality and Individual Differences*, 23, 1065-1069.
- Chaney, E. F., Roszell, D. K., & Cummings, C. (1982). Relapse in opiate addicts: a behavioral analysis. *Addictive Behaviors*, *7*, 291-297.
- Cheetham, A., Allen, N. B., Yucel, M., & Lubman, D. I. (2010). The role of affective dysregulation in drug addiction. *Clinical Psychology Review*, 30, 621-634.
- Cole, P. M., Michel, M. K., & O'Donnell, T. L. (1994). The development of emotion regulation and dysregulation: A clinical perspective. *Monographs of the Society* for Research in Child Development, 59, 73-100.
- Cooper, L., M, Frone, M., R, & Russell, M. (1995). Drinking to regulate positive and negative emotions: A motivational model of alcohol use. *Journal of Personality and Social Psychology*, *69*, 990-1005.
- Costa, P. T., & McCrae, R. R. (1984). Personality as a lifelong determinant of well-being. In C. Malatesta & C. Izard (Eds.), *Affective processes in adult development and aging*. Beverly Hills, CA: Sage.
- Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, *16*, 297-334.
- D'Zurilla, T. J., & Sheedy, C. F. (1991). Relation between social problem-solving ability and subsequent level of psychological stress in college students. *Journal of Personality & Social Psychology 61*, 841-846.
- D'Zurilla, T. J., Chang, E. C., Nottingham, E. J. I., & Lino, F. (1998). Social problemsolving deficits and hopelessness, depression, and suicidal risk in college students and psychiatric inpatients. *Journal of Clinical Psychology*, *54*, 1091-1107.
- D'Zurilla, T. J., Nezu, A. M., & Maydeu-Olivares, A. (1998). *Manual for the Social Problem-solving Inventory-Revised.* North Tonawanda, New York.: Multi-Health Systems.
- D'Zurilla, T. J., Nezu, A. M., & Maydeu-Olivares, A. (2002). *Manual for the Social Problem-solving Inventory-Revised (SPSI-R)*. North Tonawanda, New York: Multi-Health Systems.

- Darke, S., Ward, J., Hall, W., Heather, N., & Wodak, A. (1991). The Opiate Treatment Index (OTI) Manual. (Technical Report Number 11). Sydney, Australia: National Drug and Alcohol Research Centre.
- Davies, M., Stankov, L., & Roberts, R. (1998). Emotional intelligence: in search of an elusive construct. *Journal of Personality and Social Psychology*, *75*, 989-1015.
- Dawe, S., Loxton, N., Hides, L., Kavanagh, D., & Mattick, R. (2002). *Review of Diagnostic* and Screening Instruments for Alcohol and Other Drug Use and Other Psychiatric Disorders (Second Edition).
- Derogatis, L. R., Rickels, K., & Rock, A. F. (1976). The SCL-90 and the MMPI: a step in the validation of a new self-report scale. *British Journal of Psychiatry*, *128*, 280-289.
- DeVon, H., Bolock, M. E., Moyle-Wright, P., Ernst, D. M., Hayden, S. J., Lazzara, D. J., et al. (2007). A psychometric toolbox for testing validity and reliability. *Journal* of Nursing Scholarship, 39, 155-164.
- Dimitrov, D. M., & Rumrill, P. D., Jr. (2003). Pretest-posttest designs and measurement of change. *Work*, 20, 159-165.
- Dixon, W. A., & Heppner, P. P. (1987). Problem-solving appraisal, stress, hopelessness, and suicide ideation in a college population. *Journal of Counseling Psychology*, *38*, 51-56.
- Dodge, K. A. (1989). Coordinating responses to aversive stimuli: Introduction to a special section on the development of emotion regulation. *Developmental Psychology*, *25*, 339-342.
- Edwards, J., Pattison, P. E., Jackson, H. J., & Wales, R. J. (2001). Facial affect and affective prosody recognition in first-episode schizophrenia. *Schizophr Res*, 48, 235-253.
- Ehring, T., Fischer, S., Schnulle, J., Bosterling, A., & Tuschen-Caffier, B. (2008). Characteristics of emotion regulation in recovered depressed versus never depressed individuals. *Personality and Individual Differences*, 44, 1574-1584.
- Evenden, J. (1999). Impulsivity: a discussion of clinical and experimental findings. *Journal of Psychopharmacology*, *13*, 180-192.
- Feinberg, T. E., Rifkin, A., Schaffer, C., & Walker, E. (1986). Facial discrimination and emotional recognition in schizophrenia and affective disorders. *Archives of General Psychiatry*, 43, 276-279.
- Felker, B., Yazel, J. J., & Short, D. (1996). Mortality and medical comorbidity among psychiatric patients: A review. *Psychiatric Services*, 47, 1356-1363.
- Fenton, W. (2000). Depression, suicide, and suicide prevention in schizophrenia. *Suicide Life Threat Behaviors*, 30, 34-49.
- Field, A. P. (2005). Discovering Statistics Using SPSS for Windows. London: Sage.
- First, M. B., Spitzer, R. L., Gibbon, M., & Williams, J. B. W. (2001). Structured Clinical Interview for DSM-IV-TR Axis I Disorders, Research Version, Patient Edition. New York.
- Flack, W. F. J., Laird, J. D., & Cavallaro, L. A. (1999). Emotional expression and feeling in schizophrenia: Effects of expressive behavior on emotional experience. *Journal of Clinical Psychology*, 55, 1-20.
- Flett, G. L., Blankstein, K. R., & Obertynski, M. (1996). Affect intensity, coping styles, mood regulation expectancies, and depressive symptoms. *Personality and Individual Differences*, 20, 221-228.

- Floyd, F. J., & Widaman, K. F. (1995). Factor analysis in the development and refinement of clinical assessment instruments. *Psychological Assessment* 7, 286-299.
- Fox, H. C., Axelrod, S. R., Paliwal, P., Sleeper, J., & Sinha, R. (2007). Difficulties in emotion regulation and impulse control during cocaine abstinence. *Drug and Alcohol Dependence*, 89, 298-301.
- Fox, H. C., Hong, K. A., & Sinha, R. (2008). Difficulties in emotion regulation and impulse control in recently abstinent alcoholics compared with social drinkers. *Addictive Behaviors*, 33, 388-394.
- Frijda, N. H. (1988). The laws of emotion. American Psychologist, 43, 349-358.
- Furukawa, T. A., Kessler, R. C., Slade, T., & Andrews, G. (2003). The performance of the K6 and K10 screening scales for psychological distress in the Australian National Survey of Mental Health and Well-Being. *Psychological Medicine*, 33, 357-362.
- Gandek, B., & Ware Jr., J. E. (1998). Methods for validating and norming translations of Health Status Questionnaires: The IQOLA project approach. *Journal of Clinical Epidemiology*, *51*, 953-959.
- Gandek, B., WareJr, J. E., Aaronson, N. K., Alonso, J., Apolone, G., Bjorner, J., et al. (1998). Tests of Data Quality, Scaling Assumptions, and Reliability of the SF-36 in Eleven Countries: Results from the IQOLA Project. *Journal of Clinical Epidemiology*, 51, 1149-1158.
- Garnefski, N., & Kraaij, V. (2006). Relationships between cognitive emotion regulation strategies and depressive symptoms: A comparative study of five specific samples. *Personality and Individual Differences* 40, 1659-1669.
- Gold, J., Queern, C., & Bachanan, R. (1999). Battery for the Assessment of Neuropsychological Status as a screening test in schizophrenia 1: sensitivity, reliability, and validity. *American Journal of Psychiatry*, 156, 1944-1950.
- Goldberg, D., & Williams, P. (1988). A Users's Guide to the General Health Questionnaire. Windsor, Berkshire, UK: NFERNELSON Publishing Co. Ltd.
- Goldberg, L. R. (1992). The development of markers for the Big-Five factor structure. *Psychological Assessment*, *4*, 26-42.
- Goldberg, L. R., Johnson, J. A., Eber, H. W., Hogan, R., Ashton, M. C., Cloninger, C. R., et al. (2006). The International Personality Item Pool and the future of publicdomain personality measures. *Journal of Research in Personality*, 40, 84-96.
- Gow, A. J., Whiteman, M. C., Pattie, M. C., & Deary, I. J. (2005). Goldberg's 'IPIP' Big-Five factor markers: Internal consistency and concurrent validation in Scotland. *Personality and Individual Differences*, 39, 317-329.
- Gratz, K. L. (2001). Measurement of deliberate self-harm: Preliminary data on the Deliberate Self-Harm Inventory. *Journal of Psychopathology and Behavioral Assessment*, 23, 253-263.
- Gratz, K. L. (2003). Risk factors for and functions of deliberate self-harm: An empirical and conceptual review. *Clinical Psychology: Science and Practice, 10,* 192-205.
- Gratz, K. L., & Roemer, L. (2004). Multidimensional assessment of emotion regulation and dysregulation: Development, factor structure, and initial validation of the difficulties in emotion regulation scale. *Journal of Psychopathology & Behavioral Assessment*, 26, 41-54.

- Gratz, K. L., & Gunderson, J. G. (2006). Preliminary data on an acceptance-based emotion regulation group intervention for deliberate self-harm among women with borderline personality disorder. *Behavior Therapy*, *37*, 25-35.
- Gratz, K. L., Lacroce, D. M., & Gunderson, J. G. (2006). Measuring changes in symptoms relevant to Borderline Personality Disorder following short-term treatment across partial hospital and intensive outpatient levels of care. *Journal* of *Psychiatric Practice* 12, 153-159.
- Gratz, K. L., Rosenthal, M. Z., Tull, M. T., Lejuez, C. W., & Gunderson, J. G. (2006). An experimental investigation of emotion dysregulation in borderline personality disorder. *Journal of Abnormal Psychology*, 115, 850-855.
- Gratz, K. L., Tull, M. T., & Gunderson, J. G. (2008). Preliminary data on the relationship between anxiety sensitivity and borderline personality disorder: The role of experiential avoidance. *Journal of Psychiatric Research* 42, 550-559.
- Greenberg, L. S. (2002). Integrating an emotion-focused approach to treatment into psychotherapy integration. *Journal of Psychotherapy Integration*, *12*, 154-189.
- Gregorich, S. E. (2006). Do self-report instruments allow meaningful comparisons across diverse population groups? Testing measurement invariance using the confirmatory factor analysis framework. *Med Care*, 44, S78-94.
- Gross, J. J., & Levenson, R. W. (1993). Emotional suppression: Physiology, self-report, and expressive behavior. *Journal of Personality and Social Psychology*, 64, 970-986.
- Gross, J. J., & John, O. P. (1995). Facets of emotional expressivity: Three self-report factors and their correlates. *Personality and Individual Differences*, 19, 555-568.
- Gross, J. J., & Monoz, R. F. (1995). Emotion regulation and mental health. *Clinical Psychology: Science and Practice*, *2*, 151-164.
- Gross, J. J., & John, O. P. (1997). Revealing feelings: Facets of emotional expressivity in self-reports, peer ratings, and behavior. *Journal of Personality and Social Psychology*, 72, 435-448.
- Gross, J. J., & Levenson, R. W. (1997). Hiding Feelings: The acute effects of inhibiting negative and positive emotion. *Journal of Abnormal Psychology* 106, 95-103.
- Gross, J. J. (1998a). The emerging field of emotion regulation: An integrative review. *Review of General Psychology*, 2, 271-299.
- Gross, J. J. (1998b). Antecedent- and response-focused emotion regulation: Divergent consequences for experience, expression, and physiology. *Journal of Personality & Social Psychology*, 74, 224-237.
- Gross, J. J. (2002). Emotion regulation: Affective, cognitive, and social consequences. *Psychophysiology*, *39*, 281-291.
- Gross, J. J., & John, O. P. (2003). Individual differences in two emotion regulation processes: Implications for affect, relationships, and well-being. *Journal of Personality and Social Psychology*, *85*, 348-362.
- Gruber, J., & Kring, A. M. (2008). Narrating emotional events in Schizophrenia. *Journal* of Abnormal Psychology 117, 520-533.
- Gunderson, J. G. (2001). *Emotion-focused therapy: Coaching clients to work through their feelings.* Washington, DC: American Psychiatric Press.
- Gur, R. E., Kohler, C. G., Ragland, D. J., Siegel, S. J., Lesko, K., Bilker, W. B., et al. (2006). Flat Affect in Schizophrenia: Relation to emotion processing and neurocognitive measure. *Schizophrenia Bulletin*, 32, 279-287.

- Gut-Fayand, A., Dervaux, A., & Olie, J. P. (2001). Substance abuse and suicidality in schizophrenia: A common risk factor linked to impulsivity. *Psychiatry Research*, *102*, 65-72.
- Guyatt, G., Walter, S., & Norman, G. (1987). Measuring change over time: Assessing the usefulness of evaluative instruments. *Journal of Chronic Diseases*, 40, 171-178.
- Hatfield, D. R., & Ogles, B. M. (2004). The Use of Outcome Measures by Psychologists in Clinical Practice. *Professional Psychology: Research and Practice*, *35*, 485-491.
- Hayes, S. C., Wilson, K. G., Gifford, E. V., Follette, V. M., & Strosahl, K. (1996). Experiential avoidance and behavioral disorders: A functional dimensional approach to diagnosis and treatment. *Journal of Consulting & Clinical Psychology* 64, 1152-1168.
- Hayes, S. C., Strosahl, K., & Wilson, K. G. (1999). *Acceptance and commitment therapy: An experiential approach to behavior change*. New York: Guildford Press.
- Haynes, S. N., Richard, D. C., & Kubany, E. S. (1995). Content validity in psychological assessment: A functional approach to concepts and methods. *Psychological Assessment*, 7, 238-247.
- Heila, H., Isometsa, E. T., Henriksson, M. M., Heikkinen, M. E., Marttunen, M. J., & Lonnqvist, J. K. (1997). Suicide and schizophrenia: a nationwide psychological autopsy study on age- and sex-specific clinical characteristics of 92 suicide victims with schizophrenia. *American Journal of Psychiatry*, 154, 1235-1242.
- Henderson, S., Andrews, G., & Hall, W. (2000). Australia's mental health: an overview of the general population survey. *Australian and New Zealand Journal of Psychiatry*, 34, 197-205.
- Horan, W. P., Green, M. F., Kring, A. M., & Nuechterlein, K. H. (2006). Does anhedonia in schizophrenia reflect faulty memory for subjectively experienced emotions? *Journal of Abnormal Psychology August*, 115, 496-508.
- Hunt, I., M, Kapur, N., Robinson, J., Bickley, H., Flynn, S., Parsons, R., et al. (2006). Suicide in schizophrenia: Findings from a national clinical survey. *Journal of Psychiatric Practice*, 12, 139-147.
- Izard, C. E. (2002). Translating emotion theory and research into preventive interventions. *Psychological Bulletin*, 128, 796-824.
- Jablensky, A., McGrath, J., Herrman, H., Castle, D., Gereje, O., Morgan, V., et al. (1999). People living with psychotic illnesss: an Australian study 1979-1989 In Department of Health: Mental Health Branch (Ed.), National Survey of Mental Health and Wellbeing (Vol. Report 4). Canberra.
- Jacobson, N., S., & Truax, P. (1991). Clinical significance: a Statistical approach to Defining Meaningful change in Psychotherapy Research. *Journal of Consulting & Clinical Psychology*, 59, 12-19.
- Josephson, B. R., Singer, J. A., & Salovey, P. (1996). Mood regulation and memory: Repairing sad moods with happy memories. *Cognition and Emotion*, *10*, 437-444.
- Jukupcak, M., Lisak, D., & Roemer, L. (2002). The role of masculine ideology and masculine gender role stress in men's perpetration of aggression and violence in relationships. *Psychology of Men and Masculinity*, *3*, 97-106.
- Kant, G. L., D'Zurilla, T. J., & Maydeu-Olivares, A. (1997). Social problem solving as a mediator of stress-related depression and anxiety in middle-aged and elderly community residents. *Cognitive Therapy & Research*, 21, 73-96.

- Kay-Lambkin, F. J., Baker, A. L., Lewin, T. J., & Carr, V. J. (2009). Computer-based psychological treatment for comorbid depression and problematic alcohol and/or cannabis use: A randomized controlled trial of clinical efficacy. *Addiction*, 104, 378-388.
- Kay-Lambkin, F. J., Baker, A., Lewin, T., & Carr, V. (2011). Acceptability of a Clinician-Assisted Computerized Psychological Intervention for Comorbid Mental Health and Substance Use Problems: Treatment Adherence Data from a Randomized Controlled Trial. *Journal of Medical Internet Research*, 13, 1-11.
- Kay-Lambkin, F. J., Baker, A. L., Kelly, B., & Lewin, T. J. (2011). Clinical-assisted computerised versus therapist-delivered treatment for depressive and addictive disorders: a randomised controlled trial. *MJA*, 195, S44-S50.
- Kay, S., Fiszbein, A., & Opler, L. (1986). The Positive and Negative Syndrome Scale (PANSS) for schizophrenia. *Schizophrenia Bulletin*, *13*, 261-276.
- Kessler, R. C., Nelson, C. B., McGonagle, K. A., Edlund, M. J., Frank, R. G., & Leaf, P. J. (1996). The epidemiology of co-occurring addictive and mental disorders: Implications for prevention and service utilization. *American Journal of Orthopsychiatry* 66, 17-31.
- Kessler, R. C., Crum, R. M., Warner, L., Nelson, C., Schulenberg, J., & Anthony, J. (1997). Lifetime co-occurrence of DSM-III-R alcohol abuse and dependence with other psychiatric disorders in the National Comorbidity Survey. *Archives of General Psychiatry*, 54, 313-321.
- Kessler, R. C., Andrews, G., Colpe, L. J., Hiripi, E., Mroczek, D. K., Normand, S. L., et al. (2002). Short screening scales to monitor population prevalences and trends in nonspecific psychological distress. *Psychological Medicine*, 32, 959-976.
- Kessler, R. C., Barker, P. R., Colpe, L. J., Epstein, J. F., Gfroerer, J. C., Hiripi, E., et al. (2003). Screening for serious mental illness in the general population. *Archives General Psychiatry*, 60, 184-189.
- Khantzian, E. J. (1990). Self-regulation and self-medication factors in alcoholism and the addictions. In M. Galanter (Ed.), *Recent developments in alcoholism* (Vol. 8, pp. 255-271). New York: Plenum Press.
- Kline, P. (1994). An Easy Guide to Factor Analysis. London: Routledge.
- Kokkonen, M., & Pulkkinen, L. (2001). Examination of the paths between personality, current mood, its evaluation, and emotion regulation. *European Journal of Personality*, *15*, pp. 83-104.
- Koons, C. R., Robins, C. J., Tweed, J. L., Lynch, T. R., Gonzalez, A. M., Morse, J. Q., et al. (2001). Efficacy of dialectical behavior therapy in women veterans with borderline personality disorder. *Behavior Therapy*, 32, 371-390.
- Kopp, C. B. (1989). Regulation of distress and negative emotions: A developmental view. *Developmental Psychology*, 25, 343-354.
- Krause, R., Steimer-Krause, E., & Hufnagel, H. (1992). Expression and experience of affects in paranoid schizophrenia. *European Review of Applied Psychology*, 42, 131--140.
- Kring, A. M., Kerr, S. L., Smith, D. A., & Neale, J. M. (1993). Flat affect in schizophrenia does not reflect diminished subjective experience of emotion. *Journal of Abnormal Psychology*, 102, 507-517.

- Kring, A. M., Smith, D. A., & Neale, J. M. (1994). Individual differences in dispositional expressiveness: Development and validation of the Emotional Expressivity Scale. *Journal of Personality and Social Psychology 66*, 934-949.
- Kring, A. M., & Neale, J. M. (1996). Do schizophrenic patients show a disjunctive relationship among expressive, experiential, and psychophysiological components of emotion? *Journal of Abnormal Psychology*, 105, 249-257.
- Kring, A. M., & Earnst, K. S. (1999). Stability of emotional responding in schizophrenia. Behavior Therapy, 30, 373-388.
- Kring, A. M., Barrett, L. F., & Gard, D. E. (2003). On the broad applicability of the affective circumplex: Representations of affective knowledge among schizophrenia patients. *Psychological Science*, 14, 207-214.
- Kring, A. M., & Werner, K. H. (2004). Emotion regulation and psychopathology. In P. Philippot & R. S. Feldman (Eds.), *The regulation of emotion* (pp. 359-385). Mahwah, NJ: Erlbaum.
- Kring, A. M., & Sloan, D. M. (2007). The Facial Expression Coding System (FACES): Development, validation, and utility. *Psychological Assessment*, *19*, 210-224.
- Kring, A. M., & Moran, E. K. (2008). Emotional response deficits in schizophrenia: Insights from affective science. *Schizophrenia Bulletin*, *34*, 819-834.
- Lambert, M. J., & Hill, C. E. (1994). Assessing psychotherapy outcomes and processes. In M. J. Lambert (Ed.), *Bergin and Garfield's handbook of psychotherapy and behavior change* (4th ed., pp. 72-113). New York: Wiley.
- Lambert, M. J., & Hawkins, E. J. (2004). Measuring Outcome in Professional Practice: Considerations in Selecting and Using Brief Outcome Instruments. *Professional Psychology - Research & Practice*, 35, 492-499.
- Lane, R. D., Quinlan, D. M., Schwartz, G. E., Walker, P. A., & et al. (1990). The Levels of Emotional Awareness Scale: A cognitive-developmental measure of emotion. *Journal of Personality Assessment*, 55, 124-134.
- Lang, P. J., Greenwald, M. K., Bradley, M. M., & Hamm, A. O. (1993). Looking at pictures: affective, facial, visceral, and behavioral reactions. *Psychophysiology*, 30, 261-273.
- Leen-Feldner, E. W., Zvolensky, M. J., Feldner, M. T., & Lejuez, C. (2004). Behavioral inhibition: Relation to negative emotion regulation and reactivity. *Personality* and Individual Differences 36, 1235-1247.
- Linehan, M. M. (1993). *Cognitive-behavior treatment of borderline personality disorder*. New York: The Guilford Press.
- Linehan, M. M. (1994). Case consultation: A borderline dilemma. In A. L. Berman (Ed.), *Suicide and Life-Threatening Behavior* (pp. 192-198.).
- Lopes, P. N., Salovey, P., Cote, S., Beers, M., & Petty, R. E. (2005). Emotion Regulation Abilities and the Quality of Social Interaction. *Emotion* 5, 113-118.
- Loughland, C., Draganic, D., McCabe, K., Richards, J., Nasir, A., Allen, J., et al. (2010). Australian Schizophrenia Research Bank: a database of comprehensive clinical, endophenotypic and genetic data for aetiological studies of schizophrenia. *Australian and New Zealand Journal of Psychiatry*(11), 1029-1035.
- Loughland, C. M., Carr, V. J., & Lewin, T. J. (2001). The NISAD Schizophrenia research register: Why do we need a database of schizophrenia volunteers? *Australian and New Zealand Journal of Psychiatry*, 35, 660-667.

- Loughland, C. M., Williams, L. M., & Gordon, E. (2002). Visual scanpaths to positive and negative facial emotions in an outpatient schizophrenia sample. *Schizophrenia Research*, *55*, 159-170.
- Lovibond, S. H., & Lovibond, P. F. (1995). *Manual for the Depression Anxiety Stress Scales*. New South Wales: Psychology Foundation of Australia.
- Marneros, A., Deister, A., & Rohde, A. (1992). Comparison of long-term outcome of schizophrenic, affective and schizo-affective disorders *British Journal of Psychiatry*, 161.
- Martin, A., White, H., Flanagan, J., Yensel, N., & Bloomberg, T. (2011). A multisite study of association between emotion dysregulation and deliberate self-harm among substance use disorder inpatients: replication and extension. *Addictive Disorders and their Treatment*, *10*, 191-198.
- Marx, E. M., Williams, J. M. G., & Claridge, G. C. (1992). Depression and social problem solving. *Journal of Abnormal Psychology*, 101, 78-86.
- McCrae, R. R., & Costa, P. T. (2004). A contemplated revision of the NEO Five-Factor Inventory. *Personality and Individual Differences*, 36, 587-596.
- Mennin, D. S., Heimberg, R. G., Turk, C. L., & Fresco, D. M. (2005). Preliminary evidence for an emotion dysregulation model of generalized anxiety disorder. *Behaviour Research and Therapy*, 43, 1281-1310.
- Meyer, T. J., Miller, M. L., Metzger, R. L., & Borkovec, T. D. (1990). Development and Validation of the Penn State Worry Questionnaire. *Behaviour Research and Therapy*, 28, 487-495.
- Mineka, S., & Sutton, S. K. (1992). Cognitive biases and the emotional disorders. *Psychological Science*, *3*, 65-69.
- Mitchell, A. J., & Lord, O. (2010). Do deficits in cardiac care influence high mortality rates in schizophrenia? A systematic review and pooled analysis. *Journal of Psychopharmacology*, 24, 69–80.
- Nelson, H. (1982). National Adult Reading Test (NART): test manual. Windsor: NFER-Nelson.
- Newman, M. G., Zuellig, A. R., Kachin, K. E., Constantino, M. J., Przeworski, A., Erickson, T., et al. (2002). Preliminary reliability and validity of the Generalized Anxiety Disorder Questionnaire-IV: a revised self-report diagnostic measure of generalized anxiety disorder. *Behavior Therapy*, 33, 215-233.
- Nezu, A. M., & Ronan, G. F. (1985). Life stress, current problems, problem-solving and depression symptoms: A intergrative model. *Journal of Consulting & Clinical Psychology*, 53, 693-697.
- Nezu, A. M., & Carnevale, G. J. (1987). Interpersonal problem solving and coping reactions of Vietnam veterans with posttraumatic stress disorder. *Journal of Abnormal Psychology*, 96, 155-157.
- Nezu, A. M., & Perri, M. G. (1989). Social problem-solving therapy for unipolar depression: an initial dismantling investigation. *Journal of Consulting and Clinical Psychology*, *57*, 408-413.
- Nolen-Hoeksema, S., Morrow, J., & Fredrickson, B. L. (1993). Response styles and the duration of episodes of depressed mood. *Journal of Abnormal Psychology*, 102, 20-28.
- Nunnally, J. C., & Bernstein, I. H. (1994). *Psychometric theory* (3rd ed.). New York: McGraw-Hill.

- Pallant, J. (2001). SPSS Survival Manual: A step-by step guide to data analysis using SPSS, Version 10. St Leonards, Australia: Allen & Unwin.
- Patrick, D. D. (2003). Dual diagnosis: substance-related and psychiatric disorders. *Nursing Clinics of North America, 38,* 67-73.
- Penick, E., Powell, B., Nickel, E., Bingham, S., Riesenmy, K., Read, M., et al. (1994). Comorbidity of lifetime psychiatric disorder among male alcoholic patients. *Alcoholism: Clinical and Experimental Research*, 18, 1289-1293.
- Pols, H. G., & Hawks, D. V. (1992). *Is there a safe level of daily consumption of alcohol for men and women?* Canberra: Australian Government Publishing Service.
- Portnoy, D. (1999). Relatedness: Where humanistic ad psychoanalytic psychotherapy converge. *Journal of Humanistic Psychology*, *39*, 19-34.
- Priester, M. J., & Clum, G. A. (1993). Perceived problem-solving ability as a predictor of depression, hopelessness, and suicide ideation in a college population. *Journal of Counseling Psychology*, 40, 79-85.
- Putnam, K. M., Silk, Kenneth R. (2005). Emotion dysregulation and the development of borderline personality disorder. *Development and Psychopathology*, *17*, 899-925.
- Rae, A. M., Joyce, P. R., Luty, S. E., & Mulder, R. T. (2002). The effect of a history of alcohol dependence in adult major depression. *Journal of Affective Disorders*, 70, 281-290.
- Raimo, E. B., & Schuckit, M. A. (1998). Alcohol dependence and mood disorders. *Addictive Behaviors*, 23, 966-946.
- Regier, D., Farmer, M., Rae, D., Locke, B., Keith, S., Judd, S., et al. (1990). Comorbidity of mental disorders with alcohol and other drug abuse: Results from the Epidemiologic Catchment Area (ECA) study. *Journal of the American Medical Association*, 264, 2511-2518.
- Regier, D. A., Farmer, M., Rae, D., Myers, J., & et al. (1993). One-month prevalence of mental disorders in the United States and sociodemographic characteristics: The Epidemiologic Catchment Area program. *Acta Psychiatrica Scandinavica 88*, 35-47.
- Rieder, C., & Cicchetti, D. (1989). Organizational Perspective on Cognitive Control Functioning and Cognitive-Affective Balance in Maltreated Children. *Developmental Psychology*, 25, 382-393.
- Ries, R., K. (2006). Co-occurring alcohol use and mental disorders. *Journal of Clinical Psychopharmacology*, 26, S30-S36.
- Rosenthal, M., Gratz, K. L., Kosson, D. S., Cheavens, J. S., Lejuez, C., & Lynch, T. R. (2008). Borderline personality disorder and emotional responding: A review of the research literature. *Clinical Psychology Review 28*, 75-91.
- Ross, H. E., Glaser, F. B., & Germanson, T. (1988). The prevalence of psychiatric disorders in patients with alcohol and other drug problems. *Archives of General Psychiatry*, 45, 1023-1031.
- Rude, S. S., & McCarthy, C. T. (2003). Emotional functioning in depressed and depression-vulnerable college students. *Cognition & Emotion*, *17*, 799-806.
- Saha, S., Chant, D., & McGrath, J. (2007). A Systematic Review of Mortality in Schizophrenia. Is the Differential Mortality Gap Worsening Over Time? Arch Gen Psychiatry, 64 1123-1131.

- Salloum, I. M., Moss, H. B., & Daley, D. C. (1991). Substance abuse and schizophrenia: impediments to optimal care. *American Journal of Drug Alcohol Abuse*, 17, 321-336.
- Salovey, P., Mayer, J., D, Goldman, S., L, Turvey, C., & Palfai, T. P. (1995). Emotional attention, clarity, and repair: Exploring emotional intelligence using the trait meta-mood scale. In P. J. W (Ed.), *Emotion, disclosure, and health* (pp. 125-154). Washington DC: APA Books.
- Salters-Pedneault, K., Roemer, L., Tull, M. T., Rucker, L., & Mennin, D. S. (2006). Evidence of broad deficits in emotion regulation associated with chronic worry and generalized anxiety disorder. *Cognitive Therapy and Research 30*, 469-480.
- Saunders, J. B., Aasland, O. G., Babor, T. F., de le Fuente, J. R., & Grant, M. (1993). Development of the alcohol use disorders identification test (AUDIT): WHO collaborative project on the early detection of persons with harmful alcohol consumption. *Addiction*, 88, 791-804.
- Schneider, F., Gur, R., C, Koch, K., Backes, V., Amunts, K., Shah, J., N, et al. (2006). Impairment in the specificity of emotion processing in schizophrenia. *American Journal of Psychiatry*, 163, 442-447.
- Schotte, D. E., & Clum, G. A. (1982). Suicide ideation in a college population. *Journal of Consulting & Clinical Psychology*, *50*, 690-696.
- Schwartz, S. M., Trask, P. C., Shanmugham, K., & Townsend, C. O. (2004). Conducting Psychological Research in Medical Settings: Challenges, Limitations, and Recommendations for Effectiveness Research. *Professional Psychology - Research* and Practice, 35, 500-508.
- Spielberger, C. D., Gorsuch, R. L., Lushene, R., Vagg, P. R., & Jacobs, G. A. (1983). Manual for the State-Trait Anxiety Inventory (Form Y). Palo Alto, CA: Consulting Psychologists Press.
- Stallman, H. M. (2010). Psychological distress in university students: A comparison with general population data. *Australian Psychologist*, 45, 286-294.
- Staples, A., & Mohlman, J. (2012). Psychometric properties of the GAD-IV and DERS in older, community-dwelling GAD patients and controls. *Journal of Anxiety Disorders*, 26, 385-392.
- Stevens, J. (1996). *Applied multivariate statistics for the social science* (3rd ed.). Mahway, NJ: Lawrence Erlbaum.
- Stewart, S. H., & Kushner, M. G. (2001). Introduction to the Special Issue on "Anxiety Sensitivity and Addictive Behaviors". *Addictive Behaviors*, 26, 775-785.
- Strakowski, S., M, Tohen, M., Flaum, M., & Amador, Z. (1994). Substance abuse in psychotic disorders: associations with affective syndromes. *Schizophrenia Research*, 14, 73-81.
- Streiner, D. L., & Norman, G. R. (1991). *Health measurement scales: A practical guide to their development and use*. Oxford: Oxford University Press.
- Sullivan, L. E., Fiellin, D. A., & O'Connor, P. G. (2005). The prevalence and impact of alcohol problems in major depression: A systematic review. *The American Journal of Medicine*, 118, 330-341.
- Swendsen, J. D., Tennen, H., Carney, M. A., Affleck, G., Willard, A., & Hromi, A. (2000). Mood and alcohol consumption: an experience sampling test of the selfmedication hypothesis. *Journal of Abnormal Psychology*, 109, 198-204.

- Swofford, C. D., Kasckow, J. W., Scheller-Gilkey, G., & Inderbitzin, L. B. (1996). Substance use: a powerful predictor of relapse in schizophrenia. *Schizophrenia Research*, 20, 145-151.
- Teesson, M., Lynskey, M., & Dengenhardt, L. (2000). Alcohol- and drug-use disorders in Australia: implications of the National Survey of Mental Health and Wellbeing. *Australian and New Zealand Journal of Psychiatry*, *34*, 206-213.
- Thompson, R. A. (1990). Emotion and Self-regulation. *Socioemotional Development*. *Nebraska Symposium on Motivation, 36,* 367-467.
- Thompson, R. A. (1994). Emotion regulation: A Theme in Search of Definition. Monographs of the Society for Research in Child Development. In: Fox NA, ed. Monographs of the Biological and Behavioral Considerations, 59, 25-52.
- Thomsen, D. K., Mehlsen, M. Y., Viidik, A., Sommerlund, B., & Zachariae, R. (2005). Age and gender differences in negative affect: Is there a role for emotion regulation? *Personality and Individual Differences*, *38*, 1935-1946.
- Tice, D. M., Bratslavsky, E., & Baumeister, R. F. (2001). Emotional distress regulation takes precedence over impulse control: If you feel bad, do it! *Journal of Personality and Social Psychology*, 80, 53-67.
- Tobin, S. A. (1991). A comparison of psychoanalytic self psychology and Carl Roger's person-centered therapy. *Journal of Humanistic Psychology*, *31*, 9-33.
- Tomasson, K., & Vaglum, P. (1995). A nationwide representative sample of treatmentseeking alcoholics: A study of psychiatric comorbidity. *Acta Psychiatrica Scandinavica*, 92, 378-385.
- Tull, M. T., Gratz, K. L., Salters, K., & Roemer, L. (2004). The role of experiential avoidance in posttraumatic stress symptoms and symptoms of depression, anxiety, and somatization. *Journal of Nervous and Mental Disease*, 192, 754-761.
- Tull, M. T. (2006). Extending an anxiety sensitivity model of uncued panic attack frequency and symptom severity: The role of emotion dysregulation. *Cognitive Therapy and Research* 20, 177-184.
- Tull, M. T., Barrett, H. M., McMillan, E. S., & Roemer, L. (2007). A preliminary investigation of the relationship between emotion regulation difficulties and posttraumatic stress symptoms. *Behavior Therapy*, 38, 303-313.
- Tull, M. T., Jakupcak, M., Paulson, A., & Gratz, K. L. (2007). The role of emotional inexpressivity and experiential avoidance in the relationship between posttraumatic stress disorder symptom severity and aggressive behavior among men exposed to interpersonal violence. *Anxiety, Stress & Coping: An International Journal*, 20, 337-351.
- 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.
- Turpin, G. (1990). Psychophysiology and behavioral assessment: Is there scope for theoretical frameworks? In P. Martin (Ed.), Handbook of behavior therapy and psychological science: An integrative approach (pp. 348-382). New York: Pergamon Press.
- Van de Vijver, F. J. R., & Leung, K. (1997). *Methods of data analysis for cross-cultural research*. Thousand Oaks, CA: Sage.

- Verheul, R., van den Bosch, L. M. C., Koeter, M. W. J., de Ridder, M. A. J., Stijnen, T., & van den Brink, W. (2003). Dialectical behaviour therapy for women with borderline personality disorder: 12-month, randomised clinical trial in the Netherlands. *British Journal of Psychiatry*, 182, 135-140.
- Vermeersch, D. A., Lambert, M. J., & Burlingame, G. M. (2000). Outcome Questionnaire: Item Sensitivity to Change. *Journal of Personality Assessment*, 74, 242-261.
- Vermillion, J., & Pfeiffer, S. (1993). Treatment outcomes and continuous quality improvement: Two aspects of program evaluation. *Psychiatric Hospital*, 24, 9-14.
- W.H.O. (1997). *Composite International Diagnostic Interview (CIDI)*. Geneva: World Health Organization.
- Watson, D., Clark, L. A., McIntyre, C. W., & Hamaker, S. (1992). Affect, personality, and social activity. *Journal of Personality and Social Psychology*, 63, 1011-1025.
- Weinberg, A., & Klonsky, E. D. (2009). Measurement of emotion dysregulation in adolescents. *Psychological Assessment*, 21, 616-621.
- Weiting, N., & Diener, E. (2009). Personality differences in emotions: Does emotion regulation play a role? *Journal of Individual Differences*, 30, 100-106.
- Whiteside, U., Chen, E., Neighbors, C., Hunter, D., Lo, T., & Larimer, M. (2006). Difficulties regulating emotions: Do binge eaters have fewer strategies to modulate and tolerate negative affect? *Eating Behaviors*.
- Williams, A., D, Grisham, J., R, Erskine, A., & Cassedy, E. (2012). Deficits in emotion regulation associated with pathological gambling. *British Journal of Clinical Psychology*, 51, 223-238.
- Williams, K. E., Chambless, D. L., & Ahrens, A. (1997). Are emotions frightening? An extension of the fear of fear construct. *Behaviour Research and Therapy*, 35, 239-248.
- Williams, L. M., Loughland, C. M., Gordon, E., & Davidson, D. (1999). Visual scanpaths in schizophrenia: is there a deficit in face recognition? *Schizophrenia Research*, 40, 189-199.
- Wortman, P. M. (1983). Evaluation research A methodological perspective. *Annual Review of Psychology*, 34, 223–260.
- Zeman, J., & Shipman, K. (1998). Influence of social context on children's affect regulation: A functionalist perspective. *Journal of Nonverbal Behavior* 22, 141-165.

APPENDIX A-1 DERS – 36 ITEM QUESTIONNAIRE

APPENDIX A-2 INFORMATION SHEET FOR HUNTER MEDICAL RESEARCH INSTITUTE VOLUNTEERS

APPENDIX A-3 INFORMATION SHEET FOR THE NEUROSCIENCE INSTITUTE OF SCHIZOPHRENIA AND ALLIED DISORDERS GROUP

APPENDIX B-1 TIME 1: INITIAL INFORMATION SHEET AND DEMOGRAPHIC INFORMATION FOR THE COMMUNITY GROUP TIME 2: INFORMATION COVER SHEET

APPENDIX C ETHICS TABLE