UNIVERSITY OF SOUTHAMPTON

FACULTY OF SOCIAL AND HUMAN SCIENCES

Psychology

Acceptance and Commitment Therapy: Cognitive Fusion and Personality Functioning

by

Helen Bolderston

Thesis for the degree of Doctor of Philosophy

May 2013

UNIVERSITY OF SOUTHAMPTON

ABSTRACT

FACULTY OF SOCIAL AND HUMAN SCIENCES

PSYCHOLOGY

Doctor of Philosophy

ACCEPTANCE AND COMMITMENT THERAPY: COGNITIVE FUSION AND PERSONALITY FUNCTIONING

by Helen Bolderston

Personality disorders (PDs) are common, chronic, mental health problems. The majority of treatment outcome research, which has focused specifically on Borderline PD, has provided substantial empirical support for Dialectical Behaviour Therapy (DBT; Linehan, 1993), particularly in terms of self-harm reduction. Nevertheless, DBT graduates can continue to experience poor personality functioning across PD diagnostic categories, Axis I disorders, and restricted lives. Acceptance and Commitment Therapy (ACT; Hayes, Strosahl, & Wilson, 1999), might be suitable as a follow-up intervention for DBT graduates, to address their continued difficulties: to date, however, there has been little empirical investigation of its utility in relation to PD. This thesis was therefore designed to examine theoretical underpinnings of ACT relevant to the development of an ACT intervention for DBT graduates.

Study 1 tested the performance of a new self-report measure of cognitive fusion (CF), the Cognitive Fusion Questionnaire (CFQ), with a mental health sample, including individuals with PD. CF is a key ACT concept, and the CFQ proved to be a psychometrically sound measure of CF with people with mental health problems. Study 2 used cross-sectional modelling to show that CF fully mediated the relationships between two PD risk factors, negative affectivity and childhood trauma, and personality functioning in adulthood. Study 3 used the CFQ to investigate the behavioural correlates of CF. These findings strengthened the possibility that an ACT-based intervention might prove effective in improving outcomes for DBT graduates. To explore this further, Studies 4 and 5 were designed as very small-scale uncontrolled treatment development trials for this population. Study 4 suggested that ACT had a positive impact on engagement in life, but produced little improvement in psychiatric symptomology. Study 5 tested a revised protocol, which yielded more consistently positive findings, with improvements in both engagement in life and psychiatric symptoms. These findings tentatively suggest that ACT may have a role to play as a DBT follow-up intervention.

TABLE OF CONTENTS

LIST OF TABLES

Chapter III

Chapter IV

Chapter V

Chapter VI

Chapter VII

Chapter VIII

LIST OF FIGURES

Chapter II

Chapter III

Chapter IV

Chapter V

Chapter VII

Chapter VIII

DECLARATION OF AUTHORSHIP

I, **Helen Bolderston**, declare that the thesis entitled

Acceptance and Commitment Therapy: Cognitive Fusion and Personality Functioning

and the work presented in the thesis are both my own, and have been generated by me as the result of my own original research.

I confirm that:

- this work was done wholly or mainly while in candidature for a research degree at this University;
- where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated;
- where I have consulted the published work of others, this is always clearly attributed;
- where I have quoted from the source of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work;
- I have acknowledged all main sources of help;
- where the thesis is based on my own work done by myself or jointly by others, I have made clear exactly what was done by others and what I have contributed myself;
- none of this work has been published before submission

Signed ……

Date ……

ACKNOWLEDGEMENTS

There are many people I would like to acknowledge and thank. First of all, my wonderful supervisors, Professor Sue Clarke and Professor Bob Remington. I thank them for their encouragement to take this step away from the security of my clinical work into the—for me—unchartered waters of research. They have guided and challenged me, shared their expertise and knowledge, and offered me wonderful opportunities, all with great humour and warmth. Their curiosity and zest for life are inspirational, and I offer them my heartfelt thanks.

I would also like to thank all my ACT friends and colleagues for their support, encouragement and guidance. Thanks go especially to David Gillanders, Frank Bond, and Kelly Wilson. What drew me to the ACT research community originally was the combination of scientific rigour and generosity of spirit so evident there, and David, Frank, and Kelly embody these qualities wonderfully.

One of the most touching aspects of my PhD experience has been the help I have received from so many colleagues and PhD friends, both in Dorset HealthCare NHS University Foundation Trust and the University of Southampton. In particular, I would like to thank Georgie Taylor, Kirsty James, Dave Witty, Jess Kingston, Jo Lancaster, Nick Maguire, and the voluntary research assistants who helped me along the way. I also want to acknowledge and thank the patients at the Intensive Psychological Therapy Service who participated in my studies, from whom I learned so much. Their courage frequently moved me to tears.

Thank you to my family. My mum, Lola Bolderston, and my dad, Gordon Bolderston, both died long before I began this PhD journey, but their love, support, capacity to persevere with what matters, and ability to laugh at the most difficult of times, were invaluable gifts to me as I carried out this work. Sadly my brother Ian died shortly after I began this PhD. The subject of this thesis is very relevant to Ian's own struggles in life, and I want to thank him for inspiring me in my work and life.

More than anyone else, I want to thank my husband, Kim. He has been alongside me every step of this journey; supporting, encouraging, celebrating, and plying me with gin when necessary! It is no exaggeration to say that this PhD would not have been possible without his love and his unhesitating willingness to have our life together turned upside down for several years.

Finally, to my friends, who are now, along with Kim, my family. In particular, I want to thank Ruth, Debra, Joan, and Liz, for listening, sympathising, laughing, putting up with me being a research bore, and reminding me that although psychological research important, friendship is vital.

x

CHAPTER I

Personality Disorder

1.1 Diagnosis

Personality disorder (PD) is defined as: 'an enduring pattern of inner experience and behaviour that deviates markedly from the expectations of the individual's culture, is pervasive and inflexible, has an onset in adolescence or early adulthood, is stable over time, and leads to distress or impairment' (DSM-IV-TR; 2000). Within this broad definition, DSM-IV-TR identifies 10 specific PDs (and a 'personality disorder not otherwise specified' diagnostic option), with a wide range of presentations and severity, with all diagnoses implying significant emotional and interpersonal difficulty. Some diagnoses, such as BPD, involve individuals engaging in high-risk behaviours such as self-harm and suicide attempts.

DSM-IV groups these 10 PDs into three clusters, as follows:

Cluster A: Odd or eccentric disorders

Paranoid PD: characterised as "a pervasive distrust and suspiciousness of others such as their motives are interpreted as malevolent"

Schizoid PD: characterised as "a pervasive pattern of detachment from social relationships and a restricted range of expression of emotions in interpersonal settings"

Schizotypal PD: characterised as "a pervasive pattern of social and interpersonal deficits marked by acute discomfort with, and reduced capacity for, close relationships, as well as by cognitive or perceptual distortions and eccentricities of behaviour"

Cluster B: dramatic, emotional or erratic disorders

Antisocial PD: characterised as "a pervasive pattern of disregard for and a violation of the rights of others", often involving impulsivity, irresponsibility and aggressiveness.

Borderline PD: characterised as "a pervasive pattern of instability of interpersonal relationships, self-image, and affects, and marked impulsivity", often involving self-damaging urges and actions.

Histrionic PD: characterised as "a pervasive pattern of excessive emotionality and attention seeking" often including self-dramatization, suggestibility and rapidly shifting and shallow expression of emotions

Narcissistic PD: characterised by "a pervasive pattern of grandiosity (in fantasy or behaviour), need for admiration, and lack of empathy"

Cluster C: Anxious or fearful disorders

Avoidant PD: characterised as "a pervasive pattern of social inhibition, feelings of inadequacy, and hypersensitivity to negative evaluation"

Dependent PD: characterised as "a pervasive and excessive need to be taken care of, that leads to submissive and clinging behaviour and fears of separation", often including significant difficulties with making decisions, and going to excessive lengths to gain support.

Obsessive-compulsive PD: characterised as "a pervasive pattern of preoccupation with orderliness, perfectionism, and mental and interpersonal control, at the expense of flexibility, openness, and efficiency"

Two additional possible PD diagnoses, depressive PD, and passive-aggressive PD are included in an appendix to DSM-IV-TR, to encourage further research into these particular diagnoses.

There have been many criticisms of the categorical approach to PD diagnosis (e.g. Skodol & Bender, 2009), due to the high level of comorbidity amongst PD diagnoses (including across different clusters), as well as heterogeneity amongst patients with the same diagnosis. The arbitrary nature of the cut-off points for diagnosis also tends to be a target for criticism, as does the common need to use the 'personality disorder not otherwise specified' diagnosis (Verheul & Widiger, 2004).

1.1.1 *Comorbidity and Dimensionality in PDs*

Outcome trials testing interventions for PD tend to focus on a single PD diagnosis, giving the impression that the study participants only met the criteria for that PD. This is extremely misleading, as there is evidence of high levels of comorbidity across diagnostic categories. For example, Westen, Shedler, & Bradley (2006) found a mean of two diagnoses per patient in a Cluster B PD sample. McGlashan et al. (2000) found a mean of one to two additional PD diagnoses with a

sample of patients with a primary diagnosis of schizotypal, avoidant, obsessive-compulsive, or borderline PD. Not surprisingly, greater levels of comorbidity have been found to be associated with poorer quality of life (Cramer, Torgersen, & Kringlen, 2006).

It has been argued that "human personality varies continuously" (APA, 2012, p. 1), and conceptualisations of PD should reflect this, with poor personality functioning being described in terms of traits or processes that cut across diagnostic categories. Several authors have developed conceptualisations along these lines (e.g. Verheul et al., 2008; Lynch & Cheavens, 2008). DSM-V will also be based on severity ratings of dysfunctional personality traits, although one of the two dimensional assessments included will yield a category-based diagnosis.

Clearly, the conceptualisation and diagnosis of PD is currently in a state of flux. For the purposes of this thesis I will for the most part use the term 'personality disorder' (PD), to be in keeping with relevant published literature. The term 'poor personality functioning' will be used when it is particularly relevant, for example when discussing a dimension-based measure of personality functioning, or in relation to patients who have personality problems that cut across several PD diagnostic categories.

Regardless of arguments about the validity of specific categories, PD in general appears to be common in the adult population, with prevalence estimates ranging from 4% in a UK community sample (Coid, Yang, Tyrer, Roberts, & Ullrich; 2006) to 15% in a US community sample (Grant et al., 2004). The difference in prevalence estimates is thought to be due to methodological differences. It is estimated that the prevalence of PDs amongst psychiatric inpatients in the UK ranges from 36% - 67% (NIMHE, 2003). PD is thought to be under-diagnosed in both community and in-patient settings (Lamont & Brunero, 2009). According to the National Institute for Mental Health in England (NIMHE, 2003), PD diagnosis is associated with an increased likelihood to suffer from other difficulties such as substance misuse, and in general, there is an extremely high level of Axis-I disorder co-morbidity with PDs (McGlashan et al., 2000; Dolan-Sewell, Krueger, & Shea, 2001). Diagnosis of PD appears to be equal amongst men and women, although there is gender-based variation within some specific PD presentations.

Given that PDs are typically chronic conditions often experienced over decades, and given that they can affect a wide range of intra and inter-personal

experiences, they can have a hugely negative impact on the individual sufferer's life, as well as on the people close to them. Individuals with PD diagnoses are also at risk of being stigmatised, by the general public and health and social care staff (Taylor, 2010). This can lead to further isolation and poor treatment outcomes. Furthermore, these common disorders place a significant burden on society in terms of (sometimes intensive and repeated) use of health and social care services and lack of engagement in paid employment by the sufferer.

1.2 Aetiology of PD

Genetic predictors. Recent twin studies have suggested that genetic factors contribute to the risk of developing PDs. For example, Kendler et al. (2008) identified three genetic factors contributing to PD in their study based on a large, Scandinavian sample. These factors did not mirror the three-cluster structure of DSM-IV-TR. In fact the first factor identified was a broad factor, which loaded onto PDs from all three clusters, and which the authors suggested represents a general PD vulnerability factor along the lines of negative emotionality/neuroticism. The remaining two factors were much more specific, and seen by the authors as representing impulsive aggression (loading onto borderline and antisocial PDs), and inhibition/introversion (loading onto avoidant and schizoid PDs).

Livesley, Jang, and Vernon, (1998), using different methodology, identified a similar genetic factor structure, also involving a broad factor that appeared to contribute to a range of PDs, which the authors labelled emotion dysregulation or neuroticism. Heritabilities in the Kendler et al. (2008) study were described as 'modest', ranging from 20% for schizotypal PD to 41% for antisocial PD. This suggests that environmental factors have an important part to play in the development of PDs.

Environmental predictors In their multivariate twin study, Kendler et al. (2008) identified three environmental predictors of PD, which more closely matched the three-cluster structure, indicating that environmental factors may have more influence than genetic factors on the co-morbidity of PDs within clusters. The authors felt unable to label these three environmental factors due to the lack of research evidence regarding the influence of environmental risk factors on PDs in general, rather than on just one or two diagnoses such as borderline PD (BPD).

However, there are some relevant, prospective data available. For example, Johnson and colleagues used large, community sample-based, longitudinal studies to examine the impact of a range of adverse childhood experiences on the development of PDs (e.g. Johnson, Cohen, Brown, Smailes, & Bernstein; 1999; Johnson, Cohen, Chen, Kasen, & Brook, 2006). The 2006 study, examining the impact of various parental behaviours during childhood on the development of PD in adulthood, found that 10 types of parental behaviour were associated with an increased risk of developing a PD, included low parental affection or nurturing, which predicted increased risk of borderline, antisocial, avoidant, depressive, paranoid, schizoid and schizotypal PDs.

The 1999 study found that a range of childhood neglect, physical abuse, and sexual abuse experiences significantly increased the likelihood of PD diagnosis in adulthood, with some specific adverse childhood experiences increasing the risk of specific PD diagnoses. For example, neglect by parents was associated with increased risk of diagnosis of antisocial, avoidant, borderline, dependent, narcissistic, paranoid, passive-aggressive, and schizotypal PDs. More specifically, childhood sexual abuse (by parents or other adults) was associated with increased risk of BPD. Overall, a childhood experience of neglect or abuse was associated with a four-fold increase in risk of developing a PD.

Good quality research studies designed to examine genetic or environmental predictors of PD are expensive, due to the large numbers of participants involved, and the time they take to complete. There are therefore a relatively small number of relevant studies currently. It appears that both genetic and environmental factors play a role, with both broad genetic (negative emotionality) and environmental (childhood neglect) predictors having been identified, along with more specific genetic and environmental predictors. Clearly, any theoretical account of PD will need to be consistent with these findings.

There are significant gaps in the research literature regarding possible environmental predictors of PD. For example, childhood experiences such as bullying have not been tested as possible risk factors. There is some correlational data linking PD and attachment problems, but to date there have been no longitudinal studies examining the role of poor childhood attachment in the development of PDs (see Mikulincer & Shaver, 2007). Finally, there is no published data concerning the possibility of aversive experiences in adulthood increasing the risk of PD diagnosis.

1.3 Treatment of PD

Until relatively recently it was generally thought that PDs could not be treated successfully (Sperry, 1995; McMain et al., 2009). Indeed, the first RCT testing a treatment for a PD (DBT for BPD) was published in 1991 (Linehan, Armstrong, Suarez, Allmon, & Heard), indicating the paucity of research evidence guiding treatment prior to that point. Few local health and social services provided adequate treatment or care, and many health and social care professionals felt unable or unwilling to engage with individuals with PD. In the UK, this led to the publication in 2003 of the NIMHE best practice guidance paper "Personality Disorder: No longer a diagnosis of exclusion." This gave guidance on the development of services for the assessment and treatment of PD. It included examples of best practice and addressed issues of staff training and supervision. Since this time, evidence-based guidance on the treatment of specific PDs (BPD and antisocial PD) has been published (NICE, 2009; 2010).

1.3.1 *Psychosocial Treatment of PD*

The treatment of PD is a broad and complex topic. To begin with, there are 10 or 12 specific PD identified in DSM-IV-TR (2000), depending on whether depressive and passive-aggressive PD are included, or not. Many patients present with comorbid Axis I (Dolan-Sewell, Kreuger, & Shea, 2001) and/or Axis II (Zimmerman & Coryell, 1990) disorders. Prevalence of Axis II co-morbidity in research samples is not always described, making it difficult to interpret the impact of an intervention (Piper & Joyce, 2001).

Many different psychotherapeutic interventions have been applied to PDs, and different interventions, based on different theories of psychopathology and healthy psychological functioning, may have substantially different aims and treatment targets. This range might include the reduction of parasuicidal behaviours in the case of DBT (Linehan, 1993) through to a "structural change in patients' personality" (Giesen-Bloo et al., 2006, p. 649), in the cases of Schema Therapy (ST) and Transference-Focussed Therapy (TFT).

Given these factors, and taking into consideration the relatively high prevalence of PDs in the general population and their detrimental impact, it might be assumed that there is a substantial, if complex, body of research evaluating

psychosocial interventions for PD. This is not the case. Relative to the volume of research evaluating treatment for Axis-I disorders, there are few PD outcome studies, with existing research largely focussed on BPD only (Duggan, Huband, Smailagic, Ferriter, & Adams, 2007). There are a number of possible explanations for this paucity of empirical evidence, and the bias towards BPD, including the high risks often associated with BPD, the fact that people with some other PDs do not tend to seek help, the under-recognition and diagnosis of PDs in general, particularly those other than BPD, and the difficulty of conducting research with people who may variously be chaotic, contact avoidant, suspicious and so on.

This section will examine a number of psychotherapeutic interventions for PDs, limiting the focus (due to space restrictions) to interventions for which there has been at least one RCT, which is not to imply that non-RCT studies, particularly conducted early in the development of an intervention, are not important. The various therapeutic approaches are organised into two broad categories; cognitive and behavioural psychotherapies, and psychodynamic and interpersonal psychotherapies. The former will be addressed in more detail than the latter, with DBT being paid particular attention, partly because there is more empirical evidence relating to DBT than to other interventions, and partly because DBT plays a substantial role in this thesis (see Chapters VII and VIII).

The evidence base relating to psychotherapeutic interventions for antisocial PD will not be reviewed. The interventions, which are often linked with the treatment of substance misuse, and/or on an inpatient basis, are not directly relevant to this thesis. They are reviewed in the relevant NICE guidelines (2010).

1.3.2. *Behavioural and Cognitive Behavioural Psychotherapies*

This group of interventions includes *behaviour therapy* (BT)*, cognitive behaviour therapy* (CBT), and *dialectical behaviour therapy* (DBT). A summary of the main methodological features and results of all of the RCTs evaluating these approaches for PD can be found in Appendix A. Acceptance and Commitment Therapy (ACT) interventions for PD will be discussed in Chapter II. The interventions included in this section address, variously, the problematic behaviours, cognitions, schema, emotions, and action urges that are viewed as contributing to the development and maintenance of PDs. They each address the specific mechanisms of

change that are hypothesised as being important, within the theoretical context of each approach. For example, Lynch, Chapman, Rosenthal, Kuo, and Linehan (2006) view the "reduction of ineffective action tendencies linked with dysregulated emotions" as being an important mechanism of change in DBT for BPD, based on the biosocial theory of BPD. In section 1.3.3 it will be seen that in contrast, the shared hypothesised mechanism of change for psychodynamic and interpersonal interventions for PD is enactment of some form within the therapeutic relationship. It should be noted however that there is little published literature empirically testing mechanisms of change in PD interventions.

1.3.2.1 *Behaviour Therapy*

For a brief overview of behaviour therapy (BT) and the learning theories in which BT is rooted, see Chapter II. In the 1970s there was some interest in the application of behavioural principles to PD, particularly using behavioural rehearsal and reinforcement through social skills training (e.g. Argyle, Trower, & Bryant, 1974). However, the few empirical studies were severely limited methodologically. In fact, there is only one published RCT of a purely behavioural intervention for a PD to date (Alden, 1989). The trial tested behavioural interventions that had empirical support for the treatment of more general interpersonal difficulties at the time, namely *exposure* (to feared and avoided situations and experiences), and *skills training* (Stravynski & Shahar, 1983). Short-term groups based on graded exposure principles significantly improvement engagement in social activities, shyness and other relevant variables, compared to a no-treatment control group, though many participants who had shown improvement remained within the clinical range on outcome variables.

One of the study conditions involved a specific focus on intimacy, encouraging participants to move from superficial interactions to closer, more satisfying contact with others. This condition was associated with greater improvements on some outcome variables than the other conditions, but also had the highest attrition rate. This possibly highlights the conflict inherent in attempting to offer an intervention that addresses what really matters to patients; there is perhaps inevitably greater anxiety and discomfort associated with personally meaningful goals than with those that are more superficial and less personal. This theme will be revisited in the section of this thesis addressing personal values within psychotherapy

(Chapter II), and the chapters focussed on ACT treatment development for PD (Chapters VII and VIII). Unfortunately, this promising, early research has not been developed and built on directly, although behavioural principles and interventions are evident in other more common PD psychosocial interventions such as CBT and DBT.

1.3.2.2 *Cognitive Behaviour Therapy*

For a general description and overview of cognitive therapy (CT) and cognitive behaviour therapy (CBT) see Chapter II. Beck, Freeman and associates (1990) first outlined a cognitive approach to PD, in which they described an evolutionary basis to the development of such disorders. They argued that behaviours and attitudes that might once have had survival value, (for example, excessive help-seeking), are now problematic as and when they conflict with cultural norms and the requirements of current situations. The model suggests that it is the poor fit between such genetically determined strategies and the current environmental context that causes difficulties. Cognitive theory maintains that schemas develop as a way to organise information and experiences, and repeated mismatches between the individual's behaviour and environmental demands, particularly in childhood, can lead to the formation of dysfunctional schemas.

Core schema are seen as playing a central role in the development and maintenance of PDs, and are viewed as being especially rigidly-held, persistent and pervasive in the case of PD, compared to those linked to Axis I disorders. Beck and associates identified core schema associated with five central areas of human experience; love, ability, moral qualities, normality and general worth. These core schema, such as 'I am unlovable' impact which experiences are attended to, and the interpretation of those experiences. This results in conditional and control schema such as 'I must be nice to everyone', which are designed to cope with core schema. These processes of schema-influenced attentional focus and interpretation of experience also result in difficult affective experiences such as depression, behaviours such as avoidance of social interaction, as well as the core schema itself being repeatedly reinforced. According to the model, the particular form the conditional and control schema take influences the particular PD that is manifest.

The two main CBT-based intervention developments for PD are those of Davidson and colleagues (Davidson et al, 2006a; Davidson et al., 2006b), and Blum

and colleagues (Black, Blum, Pfohl, & St. John, 2004; Blum et al., 2008). Schema Therapy will be discussed in Section *1.3.3.2*. The Davidson et al. protocol involves 30 one-to-one sessions of modified CBT over a one-year period. Interventions stem from a CBT formulation of the individual's problems, and focus on reducing the negative impact of beliefs, schema and behaviours that adversely affect the participant's functioning in life. Patient and therapist jointly prioritise therapeutic goals, although high-risk behaviours such as self-harm are given priority.

The one published empirical evaluation of this intervention (Davidson et al., 2006a,b) was designed not only as a test of the intervention, but also as a means of improving on the quality of BPD therapy outcome research available at the time. There were no between-group differences found on the primary outcome measure, a composite of suicidal acts and relevant hospital visits, or indeed on the majority of secondary outcome measures. Participants appear to benefit from both conditions, though some were still self-harming at the end of the study, and still fell in the clinical range on measures such as the Beck Depression Inventory (BDI-II; Beck, Steer, & Brown, 1996).

Blum and associates (Black et al., 2004; Blum et al., 2008) reported similar outcomes from their RCT of *Systems Training for Emotional Predictability and Problem Solving* (STEPPS), a group-based, intervention for BPD that is predominantly CBT-based, with an emphasis on skills development. It includes a systems element, in that significant others in the patient's life, including healthcare professionals, are educated about BPD and are given guidance on how to most effectively interact with the patient. The approach is viewed as an adjunct to TAU (which might include one-to-one therapy), rather than as a stand-alone therapeutic intervention. The intervention is manualised and consists of 20, highly structured, weekly group sessions, and is described by the authors as easy for appropriate professionals to learn and to deliver effectively. In the one published RCT of the STEPPS approach, (Blum et al., 2008), participants in both the STEPPS and control (TAU) conditions improved on all outcome measures, but with no between-group differences in terms of self-harm, suicide attempts, or hospital admissions. It is now the authors' view that these important behavioural changes may be unlikely to occur within a 20-week intervention.

These two relatively recent studies are of better quality than much of the published PD research, in terms of sample size and other design issues, and as such,

the results can perhaps be viewed with more confidence than data from less well-designed studies. It is clear from the existing outcome research that the field of cognitive interventions for PD is in its infancy. This is surprising given the dominance of cognitive approaches as empirically supported interventions for Axis I disorders. Of course, many CBT clinicians treating patients with mood or anxiety disorders are likely to be seeing some with co-morbid PDs, and perhaps are achieving good therapeutic results. However, there is little research examining the impact of these kinds of interventions for people with PD diagnoses.

1.3.2.3 *Dialectical Behaviour Therapy*

Arguably, dialectical behaviour therapy (DBT; Linehan, 1993) is the only psychotherapy for PD that currently meets the Division 12 Task Force criteria for a well-established psychotherapy (Chambless & Hollon, 1998). It certainly has by far the most empirical support amongst treatments for PDs (Lynch, Trost, Salsman, & Linehan, 2007). However, it has only been tested to this extent as a treatment for BPD, with more limited data available in relation to other diagnoses. For this reason, the NICE guidelines (2009) currently recommend considering DBT specifically as an intervention for women with BPD who are self-harming.

1.3.2.3.1 Overview of DBT

DBT is a manualised psychotherapy originally developed to treat women meeting BPD diagnostic criteria, engaging in parasuicidal behaviours (Linehan, Armstrong, Suarez, Allmon, & Heard, 1991; Linehan, 1993a, b). It is a cognitive behaviour therapy that utilises both change and acceptance strategies, thus sharing common ground with other '3rd wave' CBTs (Hayes, 2004).

As its name suggests, DBT is essentially a behaviour therapy, and combines a range of behavioural strategies such as on-going functional analysis, behavioural experiments, contingency management, skills training, and exposure, with practices and principles rooted in dialectical philosophy and Zen Buddhism. The dialectical perspective assumes that every event or experience contains polarity, with each opposing position (referred to in dialectical philosophy as the 'thesis' and 'antithesis') being seen as valid, even if apparently oppositional and contradictory. The tension between, and the synthesis of such polarities, is seen as the process

through which change and progress can be made. The fundamental dialectic in DBT is between on the one hand, fully accepting the patient as they are, and on the other, the urgent need for them to change. This dialectical world-view permeates all of DBT, influencing every aspect of the therapy, including the basic structure of DBT, the structure and content of skills training modules, how therapist/patient conflict is addressed in the therapy, as well as specific therapeutic interventions.

DBT integrates a number of features of Zen Buddhism, both in terms of therapist assumptions and attitudes, and specific interventions and skills. Mindfulness, for example, is seen as a core skill in DBT, and is taught in a specific way that is tailored to the capabilities and needs of people with BPD diagnoses.

DBT is a comprehensive treatment package, as recommended for example by the NICE guidelines for treatment of BPD (2009). DBT interventions are designed to serve five functions; enhancing capabilities, increasing motivation, enhancing generalisation, structuring the environment, and enhancing therapist motivation and capabilities (Lynch et al., 2007). These functions can be addressed in a number of ways in DBT, but typical interventions or therapeutic modes that serve the five functions are (in turn), a skills training group, one-to-one therapy, out-of-hours telephone consultation between patient and therapist, clear DBT service leadership, and weekly DBT consultation meetings for therapists. DBT has been outlined in terms of four stages in addition to a formal pre-treatment phase, as follows:

Pre-treatment. Designed to address orientation to the therapy, motivation, and commitment.

Stage 1. Attaining basic capacities in the service of reducing high-risk behaviours.

Stage 2. Addressing Posttraumatic Stress Disorder and exposure to previously avoided emotional experiences.

Stage 3. Addressing other Axis I disorders, relationship, occupational and other issues of ‘”ordinary happiness and unhappiness” (Robins and Chapman, 2004).

Stage 4. This stage was not described in detail in the original DBT references (e.g. Linehan 1993a), but is described by Lynch et al., (2007, p. 185) as “helping the client develop the capacity for freedom and joy.” All of the published research to date relates to Stage I. This stage of treatment is described in the literature in a great deal of detail, which can only serve to guide and support clinicians, and thus perhaps contribute to reasonable adherence to the approach in non-research, more typical

clinical settings. There is little published literature delineating the features of later stages of the therapy.

1.3.2.3.2 Biosocial Model of BPD

DBT is based on a biosocial model delineating the aetiology of BPD (Linehan, 1993a), in which BPD is conceptualised as a disorder of emotion regulation, with consequent emotional difficulties negatively impacting interpersonal, behavioural, self and cognitive functioning. The biosocial model describes an on-going transactional process that can commence early in childhood, through which biologically based emotional vulnerabilities on the one hand, and environmental invalidation and consequent poor emotion-modulation on the other, interact, amplifying emotion dysregulation and its subsequent negative impact on other areas of experience.

Writing in 1993, Linehan theorised that the emotional vulnerability associated with BPD is biologically underpinned. She considered the possible role of problems in the limbic system for example, though at the time there was limited relevant empirical evidence to develop this aspect of the model. Crowell et al. (2009) have revisited the biosocial model, reviewing a range of biological, genetic, and psychosocial research, which broadly speaking, offers support to the model. Taking a developmental perspective, they also develop the model by outlining evidence that suggests that impulsivity may emerge earlier than emotional dysregulation amongst those who go on to develop BPD. Furthermore, impulsivity may contribute to the development of problems with emotional dysregulation. Finally, Crowell and colleagues outline a series of testable hypotheses related to the aetiology of BPD, directly based on the biosocial model.

This model has a number of strengths. As outlined above, it can be tested empirically, and has some empirical support. It also has clinical utility, in that it provides a non-stigmatising explanation for both staff and patients as to why people with BPD diagnoses often behave in challenging and anxiety-provoking ways. One limitation of the model is that it was originally developed to explain the aetiology of just BPD, despite evidence suggesting that there are general genetic and environmental factors that contribute to the development of many PD diagnoses, and that BPD is rarely a lone PD diagnosis. Lynch and colleagues (e.g. Lynch &

Cheavens, 2008) have however recently adapted the model to address other PD presentations (see Section 1.3.2.3.5).

1.3.2.3.3 Supporting Evidence: DBT and BPD

There is a sufficient body of DBT-related outcome research to have warranted a number of review articles (e.g. Robins & Chapman, 2004, and Lynch et al., 2007). At the time of the 2007 review, there were seven published RCTs testing DBT as a treatment for BPD. An RCT comparing DBT, TFT, and a dynamic supportive intervention, (Clarkin, Levy, Lenzenweger, & Kernberg, 2007), and a larger-scale RCT comparing DBT with manualised general psychiatric management (McMain et al, 2009) have been published since. All trials are based in community/out-patient settings (as DBT was originally designed as an out-patient treatment), although promising data from a non-randomised trial is available for an in-patient adaptation of DBT for BPD (Bohus et al., 2002).

The first RCT examining the effectiveness of DBT for BPD (Linehan et al., 1991; 1993; 1994) showed DBT performing significantly better than TAU in terms of reduction in the number of parasuicidal acts, medical risk associated with those acts, and the number of psychiatric in-patient days. However, there were no between-group differences found for depression, hopelessness or suicidal ideation, with improvements on these variables being reported for both conditions. Not all post-intervention improvements were maintained at 6 and 12-month follow-up, though the authors concluded that overall, DBT remained superior to TAU.

This general pattern of findings, that is, reduction in high risk behaviours compared to a control condition but less impact on Axis I psychopathology and other variables, has been replicated in other RCTs of DBT with high risk BPD patients (e.g., Linehan et al., 2006). Interestingly the authors of this latter study concluded that the lack of a between-group effect for these variables was not due to inadequate statistical power, as it might have been in early studies with fewer participants (e.g., Linehan et al., 1991). This issue will be discussed in more detail in Section 1.3.3.3.

One criticism that has been levelled at the DBT evidence base (Scheel, 2000; Brazier et al., 2006) is that the control interventions in the majority of published RCTs are of poor quality (see Section 1.3.2.3.6 for a more detailed discussion of this and other criticisms). Brazier et al. recommend further trials directly comparing more than one established, manualised treatment to address this issue. There has been only

one such RCT to date, comparing DBT, Transference-Focussed Psychotherapy (TFP), and Supportive Psychotherapy (SPT), conducted by Clarkin and colleagues (Clarkin, et al., 2004; 2007). Ninety patients with BPD diagnoses (history of parasuicidal behaviour was not a requirement) were randomly assigned, each participant receiving 12 months of psychotherapy. All conditions resulted in significant improvements in a number of variables including depression, anxiety, global functioning and social adjustment. DBT and TFP were equally associated with reduction in suicidality, whilst TFP and SPT were equally associated with improvements in anger. Overall, TFP was associated with changes in more variables than either DBT or SPT, although data was not reported on several important variables that DBT has been shown in the past to positively impact, such as hospitalisation rates.

1.3.2.3.4 DBT, BPD and Axis I Disorders

There is a high degree of co-morbidity between BPD and Axis I disorders such as mood disorders, anxiety disorders, eating disorders and substance misuse (Zanarini, Frankenburg, Hennen, Reich, & Silk, 2004; Harned et al., 2008). The majority of the RCTs testing the efficacy of DBT for BPD have examined its impact on such disorders. In general, where DBT has been used primarily to address high-risk behaviours, that is as a Stage 1 DBT intervention, (for example Linehan et al., 1991; McMain et al., 2009), some significant improvements in Axis I disorders such as depression, have been demonstrated. However, there tend to be few significant differences between DBT and control conditions with regards to impact on Axis I symptomology. Furthermore, with these high-risk samples, despite some improvements in Axis I symptomology, participants still often fall into the clinical range for these disorders. For example, in the McMain et al. study, depression reduced from the severe to moderate range for both conditions.

In studies where DBT has been used with lower risk BPD patients, arguably as a Stage 2/3 DBT intervention, (for example Koons et al., 2001), greater impact on Axis I disorders has been reported, with some differences between conditions in favour of DBT. In the Koons et al. study, this is the case for depression, hopelessness, and anger, where for example, 60% of the DBT condition compared to 20% of the control condition met the criterion for clinically significant change for depression.

It has been argued (Robins & Chapman, 2004; Linehan et al. 2006) that these findings make sense in terms of the DBT treatment model, in that with more risky patients, in accordance with the DBT treatment hierarchy, therapists should be targeting parasuicidal behaviours rather than Axis I disorders. However, with BPD patients who are not engaging in high-risk behaviours, therapy is more likely to be focussed on Axis I symptomology. The implication is that DBT is working as it should, regardless of the type of disorder it is addressing. However, the data do also suggest that DBT does not necessarily bring about clinically significant changes in both risky behaviours and Axis I symptomology *at the same time*. It does appear to be the case, for people with BPD diagnoses including parasuicidal behaviours, that they are likely to be far more behaviourally stable following 12 months of DBT, but they are also still likely to be experiencing considerable Axis I symptomology. To date, there is no published DBT-related research indicating how to help patients in this position move forward in their lives.

1.3.2.3.5 DBT for Conditions Other Than BPD

Adaptations of DBT for the treatment of diagnoses other than PD, such as eating disorders and substance misuse will not be examined here (for a review see Lynch et al., 2007). However, Lynch and associates have developed and are trialling an adapted form of DBT for treatment-resistant depression (TRD) accompanied by some cluster A and C PDs (Lynch, 2000; Lynch, Morse, Mendelson, & Robins, 2003; Lynch et al., 2006). Lynch notes that individuals with TRD often have co-morbid PDs, particularly Paranoid PD, Obsessive-Compulsive PD, and Avoidant PD (referred to as emotionally over-controlled PDs), and hypothesises that certain maladaptive behaviours, emotional coping strategies, and interpersonal styles associated with these PDs are likely to make treatment of TRD more difficult and less successful, and should be directly targeted in therapy for TRD.

Lynch suggests that these PDs, and indeed TRD, share common features such as psychological rigidity, risk aversion, and over-control of emotional expression (Lynch, 2000; Lynch & Cheavens, 2008). He has developed an adapted form of Linehan's (1993) biosocial model to account for the development of these personality features, hypothesising a genetic vulnerability for increased sensitivity to threat and insensitivity to reward, transacting with a childhood environment characterised by over-emphasis of performance and evaluation (Lynch & Cheavens).

This form of DBT shares many structural elements and much content with DBT for BPD, but primarily aims to "maximise openness and flexibility to new experience as well as to reduce rigid thinking and corresponding behavior" (Lynch & Cheavens, 2008, p. 166). A significant adaptation is the introduction of a new skills module addressing "Radical Openness", focussing on openness to new experience and loving-kindness/forgiveness.

Three RCTs testing versions of standard DBT with TRD have been published to date (Lynch et al., 2003; Lynch et al., 2006; Harley, Sprich, Safren, Jacobo &, Fava, 2008), and can be seen as steps towards the application of a more tailored DBT approach for TRD and PD. All three studies have shown DBT plus antidepressant medication to be superior to medication alone in the reduction of TRD symptomology. Furthermore, Lynch et al. (2006), based on an older adult TRD plus PD sample, reported that both conditions were associated with some improvements in PD symptomology, with the DBT condition in their study being associated with significantly larger reductions in interpersonal variables such as aggression and interpersonal sensitivity, which are commonly associated with PDs. The modified form of DBT outlined above is currently being trialled (Lynch et al., 2011 – 2016).

1.3.2.3.6 Criticisms of the DBT evidence base

Although there are now a number of RCTs providing empirical support for DBT as a treatment for BPD, there are weaknesses to this evidence base. For the most part, these studies involve small numbers of participants, leading to a lack of statistical power to detect treatment effects, and making it difficult to generalise from study findings (Scheel, 2000). The majority of the existing RCTs have relatively short follow-up periods (typically 6 or 12 months), when the long-term nature of BPD is considered. Westen (2000) recommended collecting 3 or even 5-year follow-up data in trials of interventions for PDs. The majority of studies have included only female participants, making it difficult to generalise from the data. The proportion of women to men with a diagnosis of BPD is 3:1 (Widiger & Frances, 1989), thus leaving the treatment of a substantial minority of BPD patients without the direct support of the data.

Writing in 2000, Scheel concluded that the quality of control groups in the DBT RCTs for BPD was problematic. Most studies have used TAU as a control condition, often resulting in a control intervention significantly less structured, less

coherent, less desirable to participants, and carried out by less experienced clinicians, than the DBT intervention. Of course, it could be argued that this is in part a result of the lack of alternative manualised psychotherapeutic interventions for BPD. However, in recent years, researchers have attempted to address this issue, by manualising and improving the quality of TAU (for example Linehan et al., 2006), and by beginning to conduct RCTs involving more than one recognised treatment (Clarkin et al. 2004; 2007).

Overall, Brazier et al. concluded in 2006 that the quality of DBT-related RCTs considered in their review was just moderate to poor, although their review was written prior to the publication of the more recent, better quality studies outlined above. In the same review, they also concluded that their findings did not indicate that DBT is a cost-effective intervention, though it might have the potential to be so in the future.

Since its development approximately two decades ago, DBT has been implemented in many health and social care settings internationally, with some arguing that the enthusiasm for the approach and the uptake of it ran ahead of the data supporting it (Westen, 2000). However, there is now a sizable body of evidence indicating that DBT is an effective intervention for reducing risk and hospitalisation for BPD patients. DBT can also be seen as having produced important positive effects in a more general sense. It has challenged the negative view within mental health services that PD, especially BPD, cannot be treated, bringing fresh hope and increased willingness to engage, for both staff and patients. This perhaps explains why DBT has been taken up with enthusiasm in many treatment settings.

Despite the positive impact of DBT, there are of course limitations to what it can achieve, as there are for all psychotherapeutic approaches. As outlined in Section 1.3.2.3.4, it is common for DBT graduates with BPD to continue to have significant problems, even after they have ceased engaging in self-harm and other risky behaviours. Many continue to experience a range of DSM-IV Axis I disorders, such as mood or anxiety disorders. Many DBT graduates also report living restricted lives, with little engagement in social or occupational activities, and with little in their lives that has meaning for them. Some continue to have a sense of self that is defined by self-harm and suicidality, despite the fact that they no longer engaging in those behaviours.

It would therefore seem important to consider the development of a psychotherapeutic intervention that could act as a secondary, post-DBT intervention. Such an intervention should be designed to enable DBT graduates to maintain progress made during DBT in terms of reduction of behavioural risk, to help them address on-going Axis-I disorders, and to offer them the possibility of building and engaging in more valued, satisfying lives. It is possible that a suitable post-DBT intervention might be found within the broad family of behavioural and cognitive behavioural psychotherapies, to which DBT itself belongs.

1.3.3 *Psychodynamic and Interpersonal Psychotherapies*

This varied group of psychotherapeutic approaches includes several forms of *psychodynamic* therapy, Cognitive Analytic Therapy (CAT; Ryle, 1990), and Schema Therapy (ST; Young, 1990). The latter, with its roots in the cognitive approach to PD could equally have been addressed with the cognitive and behavioural approaches (Section 1.2.1.1). However, it is included here as it shares important common ground with the other therapies in this section. The key unifying aspect of these approaches is that they are based on the hypothesis that therapeutic change is primarily achieved through enactments in the therapeutic relationship. For example, ST for PD has a focus on partial re-parenting (through which, according to the model, changes in maladaptive schemas are achieved), while Brief Relational Therapy (Muran, Samstag, Safran & Winston, 2005) is designed to address ruptures in the therapeutic alliance.

In their RCT comparing ST and TFT, Giesen-Bloo et al. (2006, p. 649) indicate that both of these approaches were designed to "bring about a structural change in the patients' personality", rather than just a decrease in parasuicidal behaviours. This is in contrast to many of the studies in the cognitive and behavioural group of interventions for PD, which tend to have the reduction of high-risk behaviours as primary outcome targets. It could be argued that this emphasis on impacting aspects of personality functioning, as well as risk (where appropriate), is another unifying feature of psychodynamic and interpersonal therapies for PD. See Appendix A for a summary of the principal characteristics and results for all relevant RCTs.

1.3.3.1 *CAT*

CAT (Ryle, 1990) is a time-limited (usually 16 or 24 sessions depending on the complexity of presenting problems) psychotherapy, rooted in a theoretical model that involves the integration of CBT and psychoanalytic approaches, particularly Object Relations Theory. It was developed to address the need in the UK public health system for a relatively short-term, empirically supported therapy that could be utilised in the treatment of a wide range of psychological difficulties, including more complex presentations.

Cognitive aspects of the therapy include sharing an explicit problem formulation with the patient, the setting of goals, taking a problem-solving stance, and the emphasis on a collaborative relationship between therapist and patient, in which the patient plays an active role (Denman, 2001). In terms of the analytic aspects of the therapy, Ryle has attempted to integrate into CAT a more readily understandable interpretation of core analytic concepts such as transference, countertransference, and projective identification (Ryle, 1994; 1998). Clarke et al. (2012) argue that CAT, with this central positioning of relationship within the therapy, may be in a particularly strong position to address the interpersonal difficulties with which all people with PD diagnoses, by definition, will struggle.

To date, there have been just two RCTs of CAT as an intervention for PD; Chanen et al. (2008), and Clarke et al. (2012). Both studies show some promising outcomes, with Clarke et al., for example, showing CAT out-performing TAU on several outcome measures, including PD diagnosis, with a mixed PD adult sample. In the Chanen et al. study, CAT was associated with significant improvements in a sample of adolescents with BPD symptomology, although there were no significant group differences between the CAT and control conditions.

However, two small to medium sized studies, addressing different PDs in different populations, one of which focussed on participants who did not meet full PD diagnostic criteria (Chanen et al., 2008), cannot be regarded as substantial empirical support for CAT with PD, particularly as both studies have significant methodological shortcomings, such as uncontrolled follow-up and control conditions not designed specifically to treat PD. Despite some encouraging, early results from these two studies for this relatively brief, low-intensity therapy, much more research is needed.

Clarke et al. raise the interesting possibility that CAT brings about improvements in broad PD symptomology by impacting difficulties in interpersonal relating, a process that by definition is associated with all PD categories. Larger-scale studies would be required to formally investigate such hypothesised mechanisms of change, through mediational analyses. However, Clarke et al. did report a significant group difference in favour of CAT on a measure of interpersonal problems, a result consistent with their view of the role of CAT with PDs.

1.3.3.2 *ST*

ST (1990) is rooted in Beck's original approach to PD, and integrates elements of CBT, object relations, humanistic and experiential approaches, particularly gestalt. It also has roots in attachment theory, with early adverse childhood experiences being seen as central to the development of PD pathology (Beckley, 2010). There is thus a developmental aspect to ST, with therapists taking a 'partial re-parenting' role, intended to address emotional needs (directly through the therapeutic relationship) that were not met during childhood. It was developed to address the kinds of entrenched difficulties associated with PD that more traditional CBT appears to be less effective with (Beckley, 2010).

Young has identified 18 early maladaptive schema (EMS), such as 'unrelenting standards' and 'social isolation', which are seen as having their roots in early, aversive experiences, cultural influences, and genetically-underpinned temperament. These EMS are not merely beliefs about self and others, but are more accurately seen as easily and repeatedly triggered themes that involve not just cognition, but memories, emotions and physical sensations.

Young sees a greater 'gap between cognitive and emotive change' (Young, 2004) in the treatment of PDs compared with other disorders, and thus ST places more emphasis on experiential and interpersonal interventions designed to bring about emotional change, than more traditional CBT interventions for Axis I disorders.

As is the case with CAT, to date just two ST RCTS have been published, Giesen-Bloo et al. (2006), and Farrell, Shaw, & Webber (2009). Both studies tested ST as a treatment for BPD in adults, and both showed promising results. For example, in a reasonably well-designed study comparing ST and TFT, Giesen-Bloo et

al. the authors reported significant improvements in a range of outcome variables, including parasuicidal acts, personality pathology and quality of life measures, for both conditions. ST outperformed TFT on all measures apart from quality of life. In a small-scale study with a poor quality, TAU control condition, Farrell et al. reported significant improvements on all outcome measures for ST, but none for TAU. The apparently broad impact of ST in both studies, leading to improvements in risky behaviours, PD pathology, general functioning, and quality of life, lends support to the view of both groups of authors, that ST addresses PD in a more comprehensive manor than some other psychosocial interventions for PD, such as DBT.

As is the situation with CAT however, although encouraging, these findings should be viewed with caution. The Farrell et al. study in particular is significantly flawed, with a small sample, poor control condition, and therapist adherence to ST not being independently rated. While Giesen-Bloo et al. is the better study, comparing two manualised, adherence-rated psychotherapies, it is also underpowered, and did not report follow-up data beyond the 3-year treatment period. The BPD sample for this study also appears to be somewhat less severe, in terms of parasuicidal behaviour, than is the case in a number of published DBT studies. Overall, substantial further research is required, particularly as ST remains untested with respect to the majority of PD diagnoses.

1.3.3.3 *Psychodynamic Psychotherapy*

The term *psychodynamic psychotherapy* (PP) is used to describe several interventions that share common ground but also differ in significant ways, both in terms of the theories on which they are based, and the form that the interventions themselves take. A detailed description of each of these approaches is beyond the scope of this review, but a brief outline of their shared features follows. In his 2005 review, Leichsenring uses the Gunderson and Gabbard (1996, p. 685) definition of PP as follows: “a therapy that involves careful attention to the therapist-patient interaction, with thoughtfully timed interpretation of transference and resistance embedded in a sophisticated appreciation of the therapist’s contribution to the two-person field”. Leichsenring suggests that the different models of PP can be conceptualised as operating on an “interpretive-supportive continuum” (Leichsenring,

2005, p. 844). For example, TFT (e.g. Clarkin et al., 2004), with its emphasis on "clarification, confrontation, and interpretation within the evolving transference relationship" (Clarkin et al., 2004, p.58) could be viewed as being nearer to the interpretive end of the continuum. Brief Relational Therapy (BRT; Muran, Samstag, Safran & Winston, 2005), focussing on the therapeutic alliance, and repairing ruptures in this relationship, would be considered a more supportive therapy.

Several RCTs of PPs for PDs have been reported (see Appendix A for a summary of methodological features and results), which is encouraging, as there has been a commonly held belief that PP is untested and unsupported by empirical research (for example see Shedler, 2010; Anestis, 2010). However, as a result of several different PP models existing, and researchers targeting differing PD samples, there is only one instance of more than one RCT testing a specific intervention with a particular PD patient group, making it difficult to summarise the literature and draw conclusions about the state of the empirical support for PP as an intervention for PD. The one exception is that with the publication of Bateman and Fonagy, 2009, there are now two RCTs testing mentalization-based treatment (MBT) as an intervention for BPD (the other MBT RCT being Bateman & Fonagy, 1999).

In general, results are promising, with (for example) significant between group differences on several variables including parasuicidal behaviour and psychopathology favouring MBT (Bateman & Fonagy, 1999; 2001; 2009), when compared to TAU or structured clinical management, for BPD patients. Unlike several of the studies evaluating DBT for high-risk BPD patients (for example Linehan, et al., 2006), Bateman and Fonagy report significant greater improvements in some Axis I psychopathology compared to a control condition. However, both MBT cohorts remained within the clinical range for both depression and anxiety following 18 months of treatment. Typical of several studies in this section, Clarkin et al. (2004), testing TFT for BPD, report significant improvements on many outcome measures, including measures of parasuicidal behaviours, but with few between group differences being reported in either study.

One of the relative strengths of the PP evidence base is that it includes a number of studies that have tested PPs as interventions for non-BPD PDs. For example, Svartberg, Stiles & Seltzer (2004), and Muran et al. (2005) tested short-term dynamic psychotherapy (STDP) and BRT respectively, as interventions for Cluster C PDs. In both studies, significant improvements (maintained at follow-up) were

reported on all outcome measures, although in neither study did the PP condition outperform control conditions.

As with the CAT and ST evidence bases, there are significant problems with the quality of PP-focussed RCTs. The majority of studies are underpowered, many have poor quality control conditions, and two studies (Bateman & Fonagy, 1999, and Muran, Samstag, Safran & Winston, 2005) have compromised randomisation procedures. Follow-up procedures vary from no follow-up data being reported (e.g. Gregory et al., 2008) through to relatively long-term, controlled follow-up (Svartberg, Stiles & Seltzer, 2004).

The conclusions drawn in the literature based on this small collection of studies varies considerably. At one extreme, Leichsenring and Leibing (2003) in their meta-analysis of the then available PP and CBT data for PDs conclude that both approaches are effective treatments for PD, with larger effect sizes reported for PP. Referring to this meta-analysis, Shedler (2010), in his review of the general efficacy of PP, concludes that there is good empirical support for PP for PDs. At the other extreme, Anestis (2010), in his critique of the Shedler review argues that the Leichsenring and Leibing meta-analysis is seriously flawed, and that the conclusions drawn by the authors and by Shedler are not supported by the data. Suffice it to say, despite some promising findings, as a result of the limitations of the research in terms of quality, and lack of replication, more studies are required before any PP could confidently be viewed as an efficacious treatment for any PD diagnosis.

1.3.4 *General Discussion of Psychosocial Treatment Literature*

Given the widespread negativity in clinical and academic settings about treatment outcomes for PDs until relatively recently (e.g., Sperry, 1995), it is heartening to note in excess of 20 RCTs evaluating psychosocial interventions for PD. The majority of these studies report positive outcomes, at least statistically, although whether this implies clinically meaningful changes for patients will be discussed below.

The vast majority of PD treatment research focuses on BPD, perhaps understandably, given the risks and demands associated with this patient group. There appears to be a broad pattern of outcomes from these BPD studies. Where high-risk

behaviours have been an inclusion criterion or have been very common in a sample, the interventions involved, for example, DBT and MBT, have led to significant reductions in these behaviours compared to control conditions. For the most part, although DBT has resulted in reductions in Axis I disorders such as depression, with these high-risk patients, in several studies it has not outperformed control conditions in relation to these disorders. In BPD studies where there appears to be lower frequency of recent engagement in parasuicidal behaviours, for example, Koons et al., (DBT; 2001) and Giesen-Bloo et al., (ST; TFT; 2006), interventions tend to be associated with improvements in both risk behaviours and other PD and Axis I psychopathology variables. The partial exceptions to this pattern are the two MBT trials (Bateman & Fonagy, 1999; 2001; 2009), where greater reductions in both risk behaviours *and* some Axis I psychopathology compared to the control condition, were reported, although many patients still remained within the clinical range for depression (for example), following 18-months of treatment.

It has been argued (Giesen-Bloo et al., 2006, p. 657) that "DBT and MBT are possibly optimal for a subgroup of patients with BPD who have prominent parasuicidal abnormalities, whereas SFT and TFP are meaningful for the wide range of patients with BPD". Whilst this speculation is useful in that it draws attention to the possibility of matching interventions to specific patient need, a great deal more, and better quality research is required before this question can be meaningfully addressed. For example, the Giesen-Bloo et al. study is one of just two tests of ST for BPD, both of which are underpowered.

Compared to BPD, the other PD diagnoses form a neglected majority in terms of empirical investigation (Duggan et al., 2007). Studies involving non-BPD patients tend to be trials of psychodynamic and interpersonal therapies, all of which report improvements on a range of PD and Axis I psychopathology variables, though a number (e.g. Vinnars, Barber, Noren, Gallop, & Weinryb, 2005) report no between-group differences. An exception is Clarke et al. (2012), which indicated CAT to be superior to TAU on the majority of their outcome measures. Clarke et al. is also in a minority in that the intervention was designed to be effective across PD categories, an approach it can be argued should be more common, given the levels of PD comorbidity. Although primarily addressing TRD and still under development, the work of Lynch and associates (also designed to impact transdiagnostic PD features),

may help shed light on the therapeutic needs of some patients with non-BPD PD presentations.

Overall, the quality of PD outcome research is improving. More recent studies have been based on larger samples (e.g. Doering et al., 2010), and better quality control conditions (e.g. Linehan et al., 2006). However, there are still sufficient problems with both the limited extent of the empirical literature, and the quality of the research, that in 2008, Øst concluded that DBT, the therapy for PD with the most empirical support, could not be considered an empirically supported treatment (EST). Other authors (Lynch, et al., 2007) argue that DBT is the only psychosocial intervention for BPD that does meet the EST criteria.

It is clear from this review that despite some promising outcomes from the existing studies, there are many gaps in the PD intervention empirical literature. For example, there has been virtually no investigation of mechanisms of change in interventions for PD, an essential aspect of psychotherapy development (Kazdin, 2007). At a more basic level still, there are just a handful of studies that focus on treatment of PDs beyond BPD, an issue that could be addressed in two ways. Firstly, following the general pattern in the literature to date, further trials could be conducted to evaluate interventions with each specific PD diagnosis. However, given the high level of comorbidity amongst PD diagnoses and the likely move away from a categorical conceptualisation of problematic personality functioning in DSM-V, another legitimate option would be to conduct intervention trials with mixed PD samples, using general personality functioning as a dependent variable.

Finally, even in the case of DBT, the therapy with the most substantial empirical support, further research is required. As outlined in Section 1.3.2.3.6, patients with histories of parasuicidal behaviours who benefit from DBT, often still report substantial psychological difficulties, and may be living unfulfilled lives of "quiet desperation" (Dimeff & Linehan, 2001, p. 2). There is a clinical need for the development of either an adapted form of DBT that can address risky behaviours *and* address Axis I and engagement in life problems more successfully, or an alternative, post-DBT intervention, designed to improve the lives of DBT-graduates, once they are behaviourally stable. If a post-DBT intervention were to be developed, basing it on a psychosocial intervention that shares common ground with DBT might make the

transition from one therapy to the other easier for patients. This might suggest a behavioural or cognitive behavioural intervention

CHAPTER II

Behavioural and Cognitive Behavioural Psychotherapies

2.1 First and Second Wave Behavioural and Cognitive Behavioural Psychotherapies

From the 1950s to the 1970s, a range of behavioural psychotherapy interventions was developed, based on the respondent and operant conditioning principles outlined by Pavlov, Skinner and others. At that time, psychodynamic theories and psychotherapies were dominant within psychology and psychiatry, and behaviourists such as Eysenck (1966) were critical of the lack of compelling empirical support for (typically lengthy) psychodynamic treatments of psychological difficulties. Time limited, empirically testable behavioural interventions were developed as alternatives (Clark, Beck, & Alford, 1999). It has been argued (Dougher & Hayes, 2000) that these behavioural approaches, referred to as first wave behaviour therapies (Hayes, 2004), fell into two broad camps; those with their roots in a Skinnerian, operant approach, involving behavioural analysis and direct contingency management, often focused on children and adults with developmental difficulties (e.g. Lovaas et al., 1973), and those rooted in Pavlovian learning theory, applying techniques such as systematic desensitisation and extinction to phobias, for example (e.g. Wolpe 1958; 1969). Behaviour analysis has continues to be influential in the treatment of people with developmental disorders to this day (e.g. Remington, et al., 2007). However, the application of behavioural principles to the treatment of adults with mental health problems (often referred to as 'behaviour therapy') has undergone radical change over the last three decades.

These changes in behavioural interventions for mental health problems have been influenced by a number of factors. Firstly, despite some empirical support, there were criticisms of behavioural models, including that behavioural models of depression did not satisfactorily account for cognitive and affective symptomology (Eastman, 1976). Adults with depression and anxiety appeared to respond according to their perceptions of a situation, rather than merely to environmental reinforcers (e.g. Bandura, 1969; 1971), and behaviour theory at the time was viewed as being unable to give an adequate account of the role of cognition. In fact Skinner (1957;

1969) had written extensively about verbal and rule-governed behaviour, arguing that direct operant processes apply to verbal behaviour as much as to non-verbal behaviour (Blackledge, 2007). However, neither the behavioural nor cognitive academic communities substantially developed this aspect of Skinner's work at the time and Hayes et al. (2006) suggest that Skinner was not able to give an adequate account of cognition, a view that has led to the development of *relational frame theory* (RFT; Hayes, Barnes-Holmes, & Roche, 2001. See Section 2.2.1.1).

Other criticisms were that supporting evidence for behavioural interventions often involved single case studies or was correlational in nature (Blaney, 1977), rather than being based on controlled trials. The predominant view in psychology and psychiatry was that there were limitations to the effectiveness and appropriateness of these kinds of behavioural interventions, particularly in the treatment of complex mental health problems.

A second wave of therapies (cognitive and cognitive behavioural therapies) developed against the backdrop of a shift in the experimental psychology world towards cognitive science (Farmer & Chapman, 2008). Despite this, Hayes et al., (2006), amongst others, (e.g. Teasdale & Barnard, 1993) suggest that the clinical cognitive models developed by Beck and others (Beck 1976; Ellis, 1962) were just that, clinically derived models that included cognitive mediators of behaviour, that had little in common with basic cognitive science. In fact Beck himself states that the most important influence on his development of cognitive theory and intervention for depression, his first application of a cognitive perspective, came from his own clinical experiences and his attempts to empirically test aspects of the psychodynamic therapy he was practicing at the time (Clark, et al., 1999). From this early work Beck (1967) concluded that negative judgements about self, the environment and the future (now referred to as the negative cognitive triad) appeared to play an important role in depression.

Ellis, Beck and others developed cognitive therapies designed to treat depression initially, though cognitive behaviour therapy (CBT) has now been applied to many psychological difficulties. These approaches are based on a cognitive model that emphasises the importance of an individual's perception of themselves and events in their lives, thus improving on one of the limitations of first wave behavioural theories. Beck's clinical cognitive model assumes that all human cognitive processing is biased; that is, cognitions can only ever be approximate

representations of events. Certain kinds of systematic biases in cognitive processing are implicated in the development and maintenance of psychological problems such as depression and anxiety. According to the model, meaning-making schemas develop though interaction between the individual and their environment (there are also genetic influences), and it is the interaction of these schemas with experience of events that can lead to the development of unhelpful, negative cognitions. Given the central role played by such cognitions and schemas in cognitive theory, bringing attention to these cognitive constructions, and addressing the content and accuracy of biases in them, is central to the cognitive interventions.

CBT is currently the dominant form of psychotherapy for mental health problems (Medical Research Council, 2010). There is a large body of outcome data supporting CBT-based approaches (see Butler et al., 2006 for a review). CBT is recommended in the NICE guidelines for a range of DSM-IV Axis I disorders, including depression and anxiety disorders. It is no exaggeration to say that CBT has revolutionised the treatment of common psychological problems, making time-limited, and empirically supported psychological interventions for common mental health problems widely available.

As happened with the first wave of behaviour therapies, however, second wave therapies are now also subject to criticism, both on theoretical and clinical effectiveness grounds. A number of 'dismantling studies' designed to shed light on the effective aspects of CBT have indicated that in fact the behavioural component tends to be as or more effective than full CBT in the treatment of depression (Jacobson et al., 1996; Dimidjian et al., 2007). These findings, though based on just two studies, challenge a central tenet of the cognitive model, namely that "change in beliefs is the primary mechanism of change in cognitive therapy" (Hollon & Beck, 2004, p. 453) and that therefore addressing negative cognitions will reduce depressive symptomology. Similarly, Jarrett, Vittengl, Doyle, and Clark (2007) found that changes in cognitive content in CBT for depression, were as predicted by the cognitive model, but that those changes did not predict therapeutic outcome. In fact, improvement in depressive symptomology predicted change in cognition.

In terms of treatment outcomes, there are concerns about the proportion of patients who do not respond to CBT or who relapse following CBT (e.g. Dimidjian et al., 2007). Furthermore, CBT is less effective for individuals with more complex presentations, especially those with both Axis I and Axis II disorders (Westbrook &

Kirk, 2004). Indeed, patients with these types of presentations, though clinically common, have routinely been excluded from studies testing the effectiveness of CBT (Seligman, 1995). In response to these limitations and criticisms, a third wave of therapies has been developed.

2.2 Third Wave Psychotherapies

Third wave behavioural and cognitive behavioural psychotherapies have been described by Hayes (2004) as being "particularly sensitive to the context and functions of psychological phenomena, not just their form, and thus they tend to emphasise contextual and experiential change strategies in addition to more direct and didactic ones. These treatments tend to seek the construction of broad, flexible and effective repertoires over an eliminative approach to narrowly defined problems" (p. 658). These therapies include DBT (Linehan, 1993), mindfulness-based cognitive therapy (MBCT; Teasdale et al., 2000), and ACT (Hayes et al., 1999). Fundamentally, what these approaches have in common is a recognition that it is not necessarily the content of personal experiences such as thoughts and emotions that lead to psychological difficulties, but rather the function of these private experiences and the way in which an individual relates to them. Patients are therefore enabled to develop a different kind of relationship to their private experiences, in which they can 'step back' and observe experiences such as thoughts and emotions, without them dominating the individual's behaviour.

Third wave therapeutic approaches have shown clinical promise, particularly for people with entrenched mental health problems. As outlined in Chapter I, DBT is an effective treatment for people with BPD. MBCT has been shown to be effective in preventing relapse in chronic depression (Teasdale et al., 2000). However, both therapies were originally developed to address specific clinical diagnoses and, although work is underway to test them with other diagnoses, this research is at a relatively early stage. Given that people with PD diagnoses tend to present with comorbid Axis-I and Axis-II difficulties, and often with more than one PD diagnosis, a third wave psychotherapy designed to be more generic in application might be of particular relevance to this patient group.

2.2.1 *ACT*

ACT is a third wave therapy combining behavioural change and acceptance interventions. Philosophically, it is grounded in functional contextualism (Biglan & Hayes, 1996), which has the aim of "the prediction and influence of events, with precision, scope and depth", (Hayes, 1993). As such, private experiences such as thoughts and emotions are viewed as not causing other behaviours, except under the influence of context, which suggests therefore that it is the context rather than the content of such experiences that should be addressed in the therapy. Also related to its philosophical basis, ACT emphasises workability (of, for example 'buying into' the content of a thought), rather than the truth or accuracy of that thought.

ACT is rooted in Relational Frame Theory (RFT; Hayes, Barnes-Holmes, & Roche, 2001), which implicates naturally occurring features of human language and cognition in the development of psychological suffering. Hayes, Luoma, Bond, Masuda, and Lillis, 2006, described RFT as a contextual theory of cognition. A clinical model is used to understand how processes such as cognitive fusion (CF) and experiential avoidance (EA) (see Section 2.2.1.2 for definitions) that are hypothesised to be the naturally occurring consequences of language (Hayes et al., 1999), lead to and maintain such suffering. As ACT is designed to address these processes, it should be effective in ameliorating psychological suffering across diagnostic categories. The development of this type of transdiagnostic therapeutic approach is encouraged by the Medical Research Council Review of Mental Health Research (MRC; 2010).

2.2.1.1 *RFT*

RFT (Hayes et al., 2001) is a contemporary behaviour analytic account of human language and cognition. It is a development of earlier work by Hayes and colleagues (e.g., Zettle & Hayes, 1982) on stimulus equivalence, which in turn addressed perceived limitations in Skinner's (1959) account of verbal behaviour (Gross & Fox, 2009). A detailed examination of RFT is beyond the scope of this thesis (see Hayes, et al., 2001, and Törneke, 2010 for comprehensive accounts), but there follows a summary of the essential points.

ACT is based on the view that human psychological suffering appears to be ubiquitous and that specific universal human processes are responsible for this

suffering, despite it taking many different forms. RFT implicates human language and cognition, hypothesising that one feature of human language in particular, a form of operant conditioning known as *derived relational responding*, is at the heart of both human flourishing and suffering.

Relational framing is essentially the learned capacity to arbitrarily relate any stimulus with any other stimuli, a capacity that has three main features:

1. *Mutual entailment*. This refers to the observation that when a relationship between two stimuli, (X is related to Y), is trained in a human with verbal capacity, the relationship Y is related to X is derived, without being directly trained. The form of these bidirectional relationships can vary. For example, if X is smaller than Y, then Y is larger than X, or if X is nearer than Y, then Y is further away than X.
2. *Combintorial mutual entailment*. Furthermore, if the relationships between X and Y and X and Z are trained, not only are the relationships between Y and X and Z and X derived without direct training, but the relationships between Y and Z and Z and Y are also derived, without the individual ever having experienced those two stimuli together.
3. *Transformation of stimulus functions*. Not only are the relationships between events and stimuli derived as outlined above, but if one stimulus has a particular function, for example as a reinforcer, then this function will be acquired by the other stimuli in the same relational frame. Additionally, if X is in a relationship of 'opposite' to Y, and X is a reinforcer, then not only will the stimulus function of X be transferred to Y, but Y will have the function of a punisher. This phenomenon continues to be observed when several relationships are combined through combinatorial mutual entailment.

These phenomena begin to be observed as young children start to develop language. Initially, relatively simple frames develop, such as frames of co-ordination (similarity or sameness). Later, more complex frames emerge, including frames of comparison (e.g. bigger/smaller and better/worse), and deictic frames, in which stimuli relationships are specified in terms of the perspective of the individual (e.g. I/you, here/there, now/then, mine/yours). To date, derived relational responding, also known as *relational framing*, has not been observed in other species.

Hayes, Strosahl, and Wilson (2012) argue that what RFT offers that is superior to earlier accounts of stimulus equivalence and derived stimulus relations (Sidman, 1971, for example), is that it is an attempt to provide a process-based

account of these relations (the process being relational framing), and applies this account across many types of stimulus relations, not just equivalence.

The ability to learn and apply relational framing offers the possibility of solving complex problems, of reasoning, of creativity, and of changing hostile environments; abilities that are essential to our survival and that have contributed to our becoming the dominant species on the planet. However, these very same cognitive capacities, the three features of relational framing described above, give us a "broadened interface with pain" (Torneke, 2010, p. 134), compared to non-verbal animals. Basically, we have the verbal tools to torment ourselves psychologically. For example, through the transfer of stimulus functions, we can become fearful of a stimulus even though we have never been in the presence of that stimulus, let alone had an aversive experience with it, if it is in a relational frame with a stimulus that does have those functions for us. Stimuli can be made psychologically present, including those from our pasts or imagined futures, with all that might be entailed emotionally, even though those stimuli are physically absent.

Hayes et al., (2006) argue that because the same processes that lead to psychological pain are also essential for our survival, it is not practical to develop a way to end or control relational framing. Therefore, approaches to psychological suffering that do not attempt to eliminate framing, but rather are designed to reduce its unhelpful effects, might be beneficial. Specifically, it is difficult to control, reverse or otherwise avoid relational framing, but it is (according to RFT) possible to contextually control the negative functions of relational frames that involve aversive stimuli. ACT was developed precisely for this purpose.

There is a growing body of research findings supportive of several aspects of RFT. For example, there is evidence that various types of framing with abstract tasks can be trained in young children (e.g. Barnes-Holmes, Barnes-Holmes, Smeets, Strand, & Friman, 2004). Also, transformation of stimulus functions can be demonstrated with combintorial mutual entailment (Dougher, Hamilton, Fink, & Harrington, 2007). For Hayes and collaborators it is important that ACT is grounded in RFT, which has developed out of basic science, rather than ACT being based on a clinical theory, a criticism they level at traditional CBT ((Hayes, et al., 2006; Wilson & Bolderston, 2011).

RFT has been the subject of criticism, with Palmer (2004) for example arguing that RFT experiments do not attend to covert verbal behaviour, and that some

account of relational framing could be integrated into Skinner's original (1957) analysis of verbal behaviour. McIlvane (2003) has questioned whether there is a substantive difference between RFT and earlier accounts of stimulus equivalence. Tonneau (2001) has also criticised RFT for perceived lack of precision and over-simplification. There has been a vigorous rebuttal of these criticisms (Barnes-Holmes & Hayes, 2002; Gross & Fox, 2009).

2.2.1.2 *Psychological Inflexibility Model*[1]

As described in Section 2.3.1, RFT suggests that although relational framing offers significant survival advantages for humans, it also has less helpful psychological consequences. For example, given that we can arbitrarily relate any stimuli, and given that verbal stimuli can take on the functions of other stimuli in a relational frame, we are likely to have many aversive psychological experiences that cannot be avoided simply by avoiding specific situations. Instead, we develop strategies to try to avoid private experiences such as memories and emotions, a process referred to as experiential avoidance (EA). Relational framing allows verbal events to dominate behaviour, so that we can respond to the thought 'I am unlovable' as if it is reality, perhaps by feeling sad or ashamed, and by avoiding intimacy. This process of taking verbal events to be reality is known as cognitive fusion (CF). A model of human functioning and suffering (Hayes & Strosahl, 2004), has been developed to account for the way in which EA, CF, and related processes impact psychological functioning.

The model suggests that it is psychological inflexibility (excessive verbal control over behaviour and the inability directly to contact environmental contingencies, reducing the ability to take action in a valued life-direction) that contributes to the development and maintenance of psychopathology. The processes addressed in the model are described dimensionally, accounting for what Hayes et al., (2012, p. 61) have referred to as the "continuous nature of human behaviour". As a dimensional model, it is in keeping with the dimensional approach increasingly being

[1] The psychological inflexibility model is often referred to as the 'ACT model' or the 'hexaflex model' in the ACT literature (e.g. Hayes et al., 2006). Accordingly, in this thesis these terms will be used interchangeably, acknowledging the possibility that the processes addressed in the model may have relevance beyond ACT.

taken to modelling personality functioning, the clinical issue relevant to this thesis (see Chapter I).

The ACT model can be expressed in terms of both psychological inflexibility and flexibility, the latter being hypothesised to be associated with psychological wellbeing (see Figures 2.1 and 2.2). The model outlines six inter-related psychological processes that contribute to psychological flexibility/inflexibility. Each of the six processes have been named and defined separately (e.g. Hayes, et al., 2006,) with the implication being that although they are related, each process could also be measured and manipulated individually.

Figure 2.1 *ACT Model of Psychological Inflexibility*

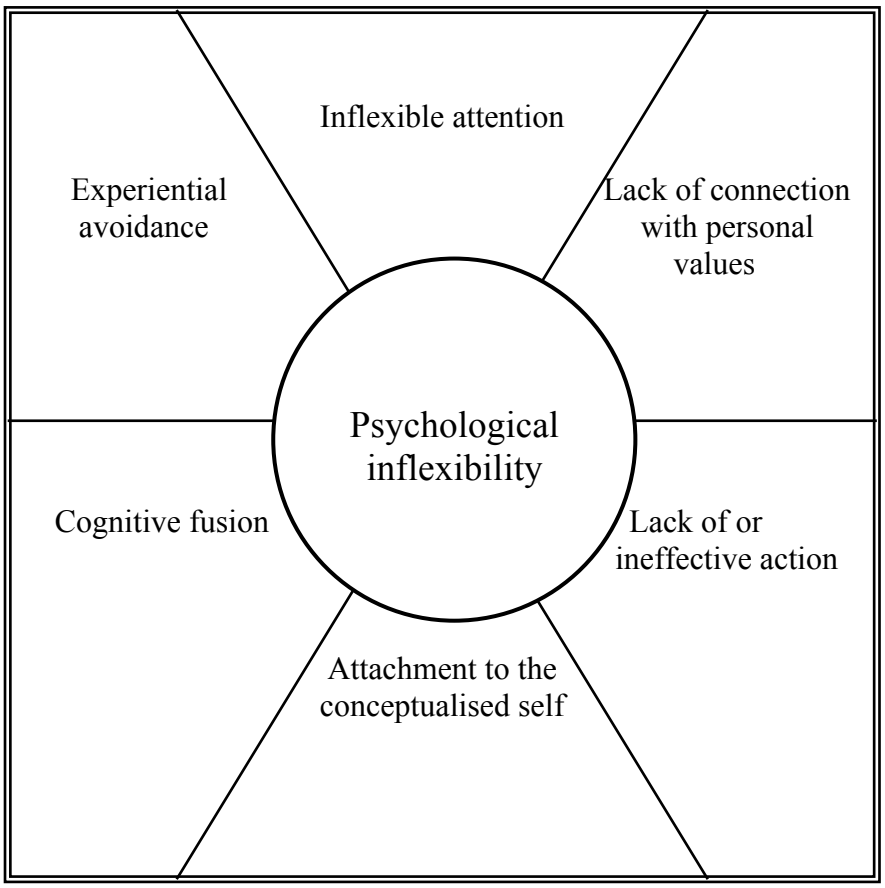
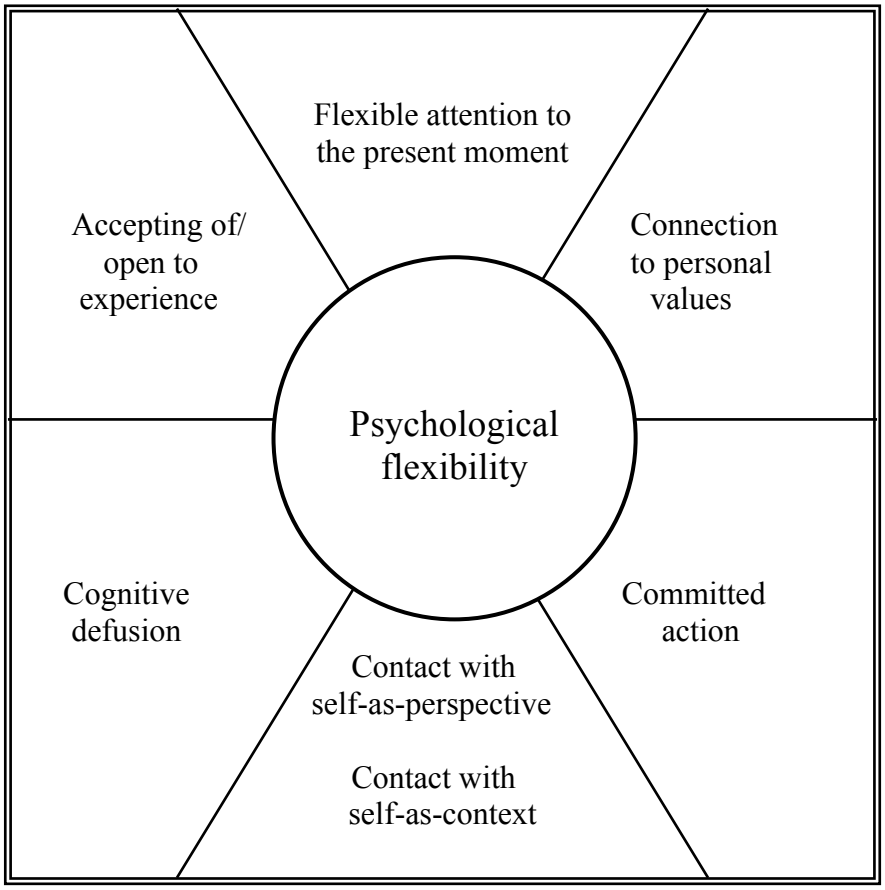


Figure 2.2 *ACT Model of Psychological Flexibility*



However, some authors have suggested that there is just a single process at work (psychological inflexibility), and that EA, CF and so on are merely facets of that process. From this perspective, "divisions we make in the model, an ostensibly unified event, are made for purely practical reasons. Breaking the model up into manageable parts allows us to talk about the whole process picture in a manageable way." (Wilson & Dufrene, 2008, p. 45). Empirical examination of the model and the processes involved would help to clarify this issue, although attempts to measure the various processes independently would be a necessary first step to facilitate such research. If standalone measures of each of the six processes were developed, this in itself would add weight to the argument that each process can be examined and manipulated independently, as well as in relation.

The components of the model such as CF, although viewed as context-determined processes in ACT theory, tend to be treated as trait-like constructs for the purposes of measurement, where individual differences are determined using self-

report measures (Cronbach & Meehl, 1955). In this thesis, the terms process and construct will both be used, depending on the context.

A number of studies in this thesis will contribute to the empirical examination of the psychological inflexibility model, both through the validation of a new self-report measure of CF, (see Chapter IV), through attempts to use this measure to examine the relationship between CF and other processes from the model, and to test the relationship between CF and poor personality functioning (see Chapters V and VI). In the meantime, in the service of clarity and a detailed understanding of the central ACT model, the six processes will each be described separately. Given that CF is the focus of this thesis, it will be addressed in greater detail than the other processes.

Inflexible Attention ---- Flexible Attention to the Present Moment

The ability to be 'present' with environmental and psychological experiences as they occur, moment by moment, is seen as beneficial in the ACT model. Being in the present moment makes it more likely that the individual will be psychologically available to be impacted by environmental contingencies, rather than merely being dominated by verbal events, and thus being in the present moment can undermine CF. At the less psychologically helpful pole of this dimension is inflexible attention, where the individual's attention is dominated by private experience such as cognitions, often relating to past experiences (memories) or possible future experiences (worries).

Examining the emotional impact of attention wandering from the present moment, Killingsworth and Gilbert (2010) conducted a large-scale study (N = 2250) using an iPhone application. They gathered data at random intervals from participants, noting what activity the participant was engaged in, where their attention was, and their current level of happiness. The results showed that mind-wandering (not being in the present moment) occurred approximately 50% of the time, was ubiquitous across activities, and was a better predictor of happiness (or lack of it) than activity. In fact mind-wandering tended to be associated with unhappiness, with time-lag analysis indicating that mind-wandering tended to precede unhappiness rather than the other way round.

Experiential Avoidance ----- Acceptance of/ Being Open to Experience

EA has been defined as “the phenomenon that occurs when a person is unwilling to remain in contact with particular private experiences (e.g., bodily sensations, emotions, thoughts, memories, behavioural predispositions) and takes steps to alter the form or frequency of these events and the contexts that occasion them” (Hayes, Wilson, Strosahl, Gifford, & Follette, 1996). Attempts to avoid private experiences are not always problematic; there are circumstances where temporary suppression of a private experience can be adaptive or at least not harmful; for example, using distraction to cope with a time-limited, one-off procedure at the dentist is unlikely to lead to lasting psychological difficulties. It is the over-reliance on and rigid application of EA that leads to and maintains psychological problems (Hayes, et al., 2012), in part at least as a result of the ‘rebound effect’ that is well documented in the thought and emotion suppression literature (e.g. Abramowitz, Tobin, & Street, 2001). In the ACT model, acceptance of or openness to experience is seen as an alternative to EA, where acceptance is defined as "actively contacting psychological experiences—directly, fully, and without needless defence” (Hayes, et al., 1996, p. 1163).

By far the largest body of empirical research connected to the inflexibility model is the large number of studies supposedly examining EA. This process has been shown to be associated with an array of psychological diagnoses, and to predict outcomes in ACT intervention studies (see Hayes, et al., 2006, for a summary of this research). The overwhelming majority of these studies have used the Acceptance and Action Questionnaire (AAQ; Hayes et al., 2004), to measure EA, but in fact the AAQ appears to measure a broader construct than EA (see Chapter VI, Section 6.1.3 for more details). The many published studies showing the relationships between EA and other variables are therefore relevant to EA, but not merely EA. In this thesis, whenever the AAQ is referred to (as opposed to the AAQII), it will be understood that the authors of the research under discussion intended to measure EA (not inflexibility), despite the measure containing some non-EA focussed items.

Testing the hypothesis that acceptance of private experiences is advantageous in terms of psychological functioning, there are a small number of component studies that have examined the impact of brief acceptance interventions, with positive results being reported. For example, Levitt, Brown, Orsillo, & Barlow (2004), compared an ACT-style acceptance intervention with thought suppression and distraction

conditions, with a sample of individuals with panic disorder. They found that the acceptance group were more willing to participate in a CO_2 challenge task, and reported feeling less anxious than the other groups, despite no physiological differences being found between the three groups.

Cognitive Fusion ----- Cognitive Defusion

CF has been defined by Strosahl, et al., (2004, p. 32) as "excessive attachment to the literal content of human thought that makes healthy psychological flexibility difficult or impossible". It has also been defined as occurring "when an individual's verbal processes (i.e., thoughts) markedly regulate overt behaviour in ineffective ways due to an inability or failure to notice the process of thinking, or context, over the products of thinking, or content" (Ciarrochi, Bilich, & Godsell, 2010, p. 53). For example, an individual might experience thoughts about a possible future event, such as the death of a loved one, as if those thoughts are descriptions of reality, or even as if those thoughts are actual current reality. This means that it is possible for an individual to be sitting next to a very much alive loved one whilst at the same time experiencing emotional and even physiological aspects of grief over their imagined future death. Aspects of fusion, as it is outlined in the ACT literature, include the dominance of thoughts in the field of awareness, and over emotion and action, the inability to view thoughts from different perspectives, psychological entanglement with thoughts, and taking the content of thoughts to be literal reality.

The risk to psychological wellbeing from CF is that when fused, the individual's behaviour is influenced more by inflexible and insensitive verbal rules, with attention being dominated by the content of thoughts, and less by direct contact with environmental contingencies. This means, for example, that an individual's behaviour may be influenced unhelpfully by a rule or other thoughts (e.g. 'I can't go out with the other students after class because I get too anxious'; 'they think I'm boring'). People with their attention caught up in these kinds of thoughts are less 'available' to register and be impacted by direct environmental contingencies, such as classmates smiling at them or momentary feelings of warmth when someone talks to them in class. The result may be increased avoidance of social situations and ultimately even the development of a self-maintaining social anxiety problem.

In the ACT literature, due to the impact of relational framing, CF along with EA, is seen as a particularly powerful determinant of psychological inflexibility, and hence these processes "prolong suffering" (Ciarrochi, et al., 2010, p. 53), and play a

central role in the development and maintenance of psychopathology Hayes & Strosahl, 2004). Furthermore, CF is viewed (Hayes, et al., 2006), as leading to and supporting EA. As an example, individuals who are fused with thoughts such as 'I am unlovable' may experience these thoughts to be reality, with their attention dominated by such thoughts, which understandably would be aversive. This might lead to EA to try and reduce the aversive impact of the thoughts. Individuals who have similar thoughts but who are able to see them as 'just thoughts' and who can view themselves from other perspectives too, are less likely to experience such aversive reactions and are therefore less likely to resort to EA.

This latter capacity is referred to as cognitive defusion, and is at the more psychologically healthy pole of this dimension. There are dozens of cognitive defusion techniques used within ACT (see Hayes and Strosahl, 2004, for examples), which have been designed to reduce fusion. There have been a number of experimental studies designed to test brief defusion interventions, usually in relation to aversive stimuli such as negative self-statements. For example, Masuda and colleagues have conducted a series of studies investigating the impact of rapid repetition of self-referential negative words (Masuda, Hayes, Sackett, & Twohig, 2004; Masuda, Hayes, Twohig, Drossel, Lillis, & Washio, 2009; Masuda, Twohig, Stormoa, Feinstein, Chou, & Wendell, 2010). De Young, Lavender, Washington, Looby, and Anderson (2010) also tested this form of defusion intervention. Watson, Burley and Purdon (2010) tested the same defusion practice in relation to contamination-focused obsessional thoughts. In contrast, Healy, Barnes-Holmes, Barnes-Holmes, Keogh, Luciano, and Wilson, (2008) tested the impact of adding the phrase "I am having the thought that" as a prefix to negative and positive self-statements. Each of these studies has shown the significant impact of a defusion practice on at least one dependent variable such as emotional discomfort, willingness to read and think about personal statements, and the believability of the target words or statements.

However, all of these studies have employed measures of believability of words and thoughts with the assumption that this is the equivalent of measuring CF, despite believability being a narrow operationalisation of CF (see Chapter IV for a discussion of this issue). Furthermore, none of the measures of believability used were established as being psychometrically sound. For example (and typical of this group of studies), Healy et al. (2008) asked participants to rate on a scale from 0 to

100 the extent to which they found each statement believable. Caution therefore should be exercised when drawing conclusions about the impact of defusion practices from this research.

There have been few correlational or mediational studies examining CF in relation to other variables. As outlined in Section 2.3.2.2, the development of the AAQ facilitated a great deal of this type of research in relation to EA/psychological inflexibility. As there has been no well-designed, broadly applicable measure of CF available until very recently, there is currently no equivalent large body of CF-focused research (see Chapter IV for a detailed examination of CF measurement). In one exception, (Zettle, Rains & Hayes, 2011), believability of thoughts (used as a proxy for CF) was found to have mediated the impact of an ACT intervention for depression. In another outcome trial (ACT for psychosis), believability of hallucinations was found to mediate the relationship between hallucination frequency and associated distress (Gaudiano & Herbert, 2006).

To summarise, there is a small amount of CF-relevant research, in the form of correlational and mediational studies, as well as laboratory-based component studies testing the impact of defusion techniques. In general, this research has yielded results supportive of the ACT model, but the lack of a good measure of CF has hindered examination of this important process.

Attachment to Conceptualised Self ----- Contact with Self-As-Perspective

In the ACT model, particular attention is paid to three experiences of self (Stewart, Villatte, & McHugh, 2012); the conceptualised self, which is the verbally constructed view of self, self as a process of moment-by-moment awareness, and self-as-perspective (also referred to as self-as-context.) Fusion with the conceptualised self is seen as leading to EA and contributing to psychological inflexibility. For example, taking a thought about oneself such as ‘I’m the kind of person who always gives up on things’ to be literal, is likely to impact on the ability to take effective action in life.

Hayes, et al., (2006, p. 9) describe this perspective on self as “the context for verbal knowing” rather than “the content of that knowing”, and when we experience ourselves as self-as-context, we experience ourselves as being not merely the content of our thoughts and feelings, but as the conscious container for them. This experience of self can be particularly helpful for individuals who are fused with negative self-

referential thoughts, in that they can have moments of experiencing themselves as not merely the content of such thoughts.

The impact on psychological health of this process as it is operationalised in the ACT model has not been empirically tested in isolation from other aspects of ACT, although exercises and metaphors designed to support experience of self-as-perspective are generally included in efficacious ACT treatment protocols. Additionally, mindfulness practice is seen as a context in which this sense of self is supported and experienced, and there is a large body of literature attesting to the psychological and physical health benefits of regular mindfulness practice (see Grossman, Niemann, Schmidt, & Walach, 2004, and Piet & Hougaard, 2011, for two meta-analyses of the clinical impact of mindfulness).

Lack of Connection with Personal Values ------ Connection with Personal Values

All cognitive and behavioural psychotherapies place some emphasis on setting behavioural goals, but ACT is unusual in its emphasis on the individual connecting with their personal values as a means of ensuring that actions taken are in the service of what is personally important to them, and to motivate engagement in specific behaviours, even over long periods of time. Values have been defined as "freely chosen, verbally constructed consequences of on-going, dynamic, evolving patterns of activity, which establish predominant reinforcers for that activity that are intrinsic in engagement in the valued behavioural pattern itself" (Wilson & Dufrene, 2008, p. 64). Less technically, they are "our heart's deepest desires: how we want to be, what we want to stand for, and how we want to relate to the world around us", and "leading principles that can guide us and motivated us as we move through life" (Harris, 2007, p. 167).

Whilst values are an important aspect of the ACT model, and there are many exercises in ACT protocols specifically designed to help individuals connect with their personal values, relatively little research has been conducted that specifically examines the role of valuing. As with several of the ACT model processes, this is at least in part due to issues of measurement. One exception is McCracken and Yang (2006), who developed a measure of values and valued action specifically for patients dealing with chronic physical pain. They found that engagement with values made a unique contribution to the prediction of patient functioning (both physical and psychosocial), independent of the predictive power of pain acceptance. Using a different approach, in a study that investigated valuing by testing a stand-alone values

intervention, Creswell, Welch, Taylor, Sherman, Gruenewald, & Mann (2005) found that in a laboratory-based experiment with a student sample, affirmation of personal values acted as a protective factor in terms of hormonal and psychological responses to a stressful task.

These two studies suggest that it is possible to examine valuing and its impact on psychological functioning, and furthermore, that as the model predicts, connection or lack of with personal values may well impact functioning, but a great deal more research is needed in this area.

Lack of or Ineffective Action ----- Committed Action

ACT is a behavioural psychotherapy, and as such the ability to engage in values-based activity, and to make values-consistent behavioural changes, is an essential aspect of the therapy. The risk according to the hexflex model, when CF and EA are predominant, and when there is a lack of connection with values, is that behaviour will be inflexible and ineffective. This may take the form of avoidance and lack of persistence, or alternatively, excessive and impulsive actions (Hayes, et al., 2012). In the hexaflex model, what is thought to contribute to psychological flexibility and thus to psychological wellbeing is "the development of larger and larger patterns of effective action linked to chosen values" (Hayes, et al., 2006, p. 9). This development of a broader behavioural repertoire is seen as allowing greater possibilities for positive reinforcement from environmental contingencies.

There is a great deal of research attesting to the negative impact of behavioural avoidance in relation to many psychological difficulties, including anxiety disorders and depression (see Barlow, 2002, and Ottenbreit & Dobson, 2004, respectively, for summaries of this literature). Similarly, impulsive and excessive behaviours such as self-harm and substance misuse, have also been linked with poor psychological functioning (e.g. Kingston, Clarke, & Remington, 2010). Interventions that encourage effective and flexible action such as behavioural activation treatments for depression, and exposure-based treatments for anxiety disorders have proved to be effective (e.g. Dimidjian et al., 2006; Ost, Thulin, & Ramnero, 2004). Although none of this research is specific to ACT, behavioural interventions in the context of ACT protocols closely resemble those developed by the wider behavioural and cognitive behavioural research community, albeit with the emphasis on behaviour being in the service of personal values.

2.2.1.2.1 *Summary*

Some efforts have been made to empirically examine the six core processes of the ACT model, in particular in relation to psychopathology, with reported findings generally supporting the model. However, there are significant gaps in the literature, particularly when considering the relationships amongst the processes, and in fact Hayes et al, in their 2006 review paper stated that at that stage, correlational study-based investigations of the model had not generally involved examining the role of the processes individually. To date there are virtually no published studies testing the ACT model in relation to complex psychological difficulties such as PD.

2.2.1.3 *ACT: Therapeutic Features*

Hayes et al., (2012) suggest that the value of a good clinical model is that it should directly indicate what would constitute effective clinical interventions. In the case of ACT, the model indicates six processes that contribute to psychological inflexibility and are thus implicated in psychopathology. The therapy was therefore developed specifically to address these processes. For a detailed description of ACT, see Hayes and Strosahl, 2004, and Hayes et al. For an outline of an ACT protocol specifically for PD, see Chapters VII and VIII. Below is a brief summary of the key features of the therapy.

As can be understood from the hexaflex model, the primary aim of ACT is to increase psychological flexibility, rather than to reduce symptomology, although the latter is often a consequence of ACT interventions, as demonstrated in many ACT outcome trials. The intention is that regardless of specific diagnosis or symptomology, excessive verbal control over behaviour will be disrupted and the capacity to make direct contact with environmental contingencies will increase, along with the ability to take action in valued life directions. Although some didactic teaching can occur in ACT, much greater emphasis is placed on setting up contexts, both within and between sessions, where the patient can become directly aware of their experiences of, for example, the impact of CF and EA. Such experiential exercises are often based on metaphors and paradoxes, which Hayes et al., (2004, p. 6) refer to as “non-linear forms of language”. Essentially, the aim of such

interventions is to bring awareness to the processes of language and cognition, rather than to provide the patient with a new set of beliefs.

It is common for overviews of ACT to describe the course of therapy in linear terms (e.g. Hayes & Smith, 2005; Ruiz, 2010), as if there is just one effective order in which to address the six core processes, and it can make theoretical and clinical sense to address some aspects of the hexaflex before others. One example of this is that supporting the cultivation of willingness and ability to experience uncomfortable thoughts and emotions is likely to be helpful before inviting an individual to take new behavioural steps in a valued direction. However, this therapeutic work rarely happens in a simple, linear fashion. One way to think about the hexaflex model is that the four processes on the left of the model, present moment awareness, acceptance, cognitive defusion, and self-as-perspective, are mindfulness-related processes, with the two processes on the right of the model, values and committed action, being concerned with behavioural change. An ACT therapist who is paying attention to the moment-by-moment experiences of the patient will be likely to move back and forth, addressing the processes on either side of the hexaflex as required. For example, when the patient begins to get fused and avoidant when focusing on values or committed action, a shift is made to address the mindfulness processes, in order to help the individual hold the thoughts and feelings that have arisen, in defused, accepting awareness. This approach mirrors that of another 3rd wave therapy, DBT, where the therapist moves flexibly back and forth between acceptance and change therapeutic strategies.

ACT handbooks have also tended to include long lists of specific experiential exercises and metaphors, to be used to address particular processes such as CF and EA (e.g. Hayes & Strosahl, 2004). As a result of this, and the linear description of ACT, it could be assumed that the therapy consists of a set of prescribed techniques that should be applied in a relatively inflexible order and manner. In fact, because ACT was developed to address the hexaflex processes, any experiential exercises, metaphors, and interpersonal exchanges between therapist and patient that facilitate awareness and direct experience of the impact of those processes, can be incorporated in the therapy, within the limits of what is safe and acceptable to the patient. Thus, some ACT therapists use a great deal of pre-planned experiential exercises, whilst others rarely use these methods, instead using more informal, conversational

interventions, informed by a moment-by-moment formulation in terms of the hexaflex processes.

This issue of rigid adherence to a therapeutic plan versus flexible responding in the moment is of importance when considering how to best test a therapy in a research outcome trial. There is a tension between the need to describe the intervention as clearly as possible and to reduce variation in participant experience on the one hand, and on the other, to represent the therapy realistically, including the capacity of the therapist to respond flexibly, which might be considered as contributing to the effectiveness of intervention. A validated ACT adherence and competence measure (Forman, Herbert, Moitra, Yeomans, & Geller, 2007), reflects this tension in that it assesses the extent to which therapists address specific processes and issues within therapy, but does not require therapists to address them using particular interventions. This tension in relation to ACT is touched on again in this thesis, in the chapters dealing with ACT treatment development trials.

With these caveats in mind, a typical order of addressing key processes and issues in ACT is as follows:

- Creative hopelessness – inviting the individual to reflect on how effective or otherwise their (typically avoidant) coping strategies have been to date
- Underlining the likelihood that automatic compulsive avoidance of private experiences such as emotions, whilst understandable, is likely to be ineffective, particularly in the long-term
- The possibility of acceptance as an alternative to avoidance
- Relationship to thoughts – cognitive fusion and defusion
- Relationship to self – conceptualised self and self-as-context
- Values
- Committed action

2.2.1.4 *Evidence Supporting ACT: Outcome and Process Findings*

Ruiz (2010) identified more than 30 RCTs where an ACT-based intervention had out-performed a control condition. These outcome studies included protocols addressing Axis-I disorders such as depression (Zettle & Hayes, 1986), psychosis

(Bach & Hayes, 2002), and social phobia (Block, 2002), as well as other conditions such as work stress (Bond & Bunce, 2000) and difficulties associated with chronic pain (McCracken et al., 2005). As with any therapeutic approach in its infancy however, there have been criticisms levelled at the quality of this evidence. Ost (2008) conducted a systematic review and meta-analysis of the ACT RCTs available at the time where ACT had been compared to CBT, concluding (based on the quality of the trials and the efficacy of the interventions), that at that point ACT could not be considered an empirically supported intervention. His methodology was criticised and his findings were reanalysed (Gaudiano, 2009), with Gaudiano refuting Ost's conclusions. Powers, Zum, Vorde, Sive, Vording, and Emmelkamp (2009) also carried out a meta-analysis, concluding that ACT outperformed waiting list control conditions, but not other established interventions. Again, these findings were reanalysed (Levin & Hayes, 2009), who, based on their corrections to the data and methods used by Powers et al., concluded that ACT did significantly outperform established interventions.

Some efforts have been made not only to test the efficacy of ACT interventions, but also to ascertain whether therapeutic gains occur through theory-consistent means. The findings of these kinds of studies have tended to support the ACT clinical model. For example, Zettle et al., (2011), in their reanalysis of data from an ACT and CBT for depression RCT (Zettle & Rains, 1989), found that post-treatment CF level mediated reduction in depression at follow-up, in the ACT condition only.

Despite the promising outcome and process data summarised above, Ost was justified in some of his criticisms of the ACT evidence base. Published outcome trials have varied considerably in size and quality, with some early trials in particular being relatively poorly designed. In fact, a similar situation with regards to research quality can be found in the early, published trials of other mental health interventions such as DBT and MBSR, perhaps suggesting that this is a common feature of psychosocial treatment development in the 1980s and 1990s. Recent improvements in study quality may be a function of improvements in funding for ACT trials, and may also be linked to the general shift towards better quality outcome trials in mental health research over the last decade or so, linked with the introduction of guidelines for conducting and reporting such studies (e.g. Boutron et al., 2008). One of the difficulties with this particular body of evidence is that because ACT was designed to be a transdiagnostic,

broadly applicable approach, it is being tested in many settings in relation to many difficulties, both psychological and physical. This means that studies are currently somewhat ‘thinly spread’.

2.2.1.4.1 ACT, Complex Mental Health Problems, including PD

There are a handful of studies suggesting that ACT might have potential with more complex or difficult-to-treat presentations including poly-substance abuse (Hayes et al., 2004), and psychotic depression (Gaudiano et al, 2007). Unlike DBT and MBCT, which were each developed for use with just one clinical group, ACT is a trans-diagnostic approach. Therefore, it might be of particular relevance to a group of patients who tend to have co-morbid and heterogeneous presentations, such as PD patients. To date only three small-scale outcome trials have tested an intervention that includes aspects of ACT, with people with PD diagnoses.

Holmes and colleagues (Holmes, Georgescu, & Liles, 2006; Hurley & Holmes, 2010) are testing a group-based psychotherapeutic intervention for BPD. The intervention, contextually-based DBT, is essentially full-programme DBT with all skills and teaching that refer to changing the content of private experiences modified in line with ACT principles. Although a trial of this intervention has not, to date, been published, the researchers have presented some interim outcome data at a conference (Hurley & Holmes). Data from 33 participants who completed the intervention indicated group improvements in depression, anxiety, and psychological flexibility. A predicted improvement in quality of life was not found, with the researchers speculating that it “is not until Stage 3 [DBT] that one is actively engaged in the process of improving quality of life.” No data on changes in self-harm or other BPD-relevant variables was presented.

The study was small-scaled and uncontrolled, both appropriate research design features for a preliminary treatment development trial. However, given these limitations, the lack of data on self-harm and suicide attempts when these are specific and important targets of the treatment, and the fact that the study has not been published in a peer-reviewed journal thus far, the promising outcomes have to be viewed with caution. In addition, the contribution of ACT to this intervention has been clearly delineated (Holmes et al., 2006), and it is evident that the intervention is predominately DBT in nature. It is consequently difficult to argue that this is really a test of ACT in relation to PD.

Gratz and Gunderson (2006) presented preliminary data on the impact of a 14-week, outpatient, group intervention for women with BPD diagnoses who were self-harming. The intervention was described as an acceptance-based emotion regulation group, emphasising "the control of behavior when emotions are present, rather than the control of emotions themselves", (p. 26). The intervention combined elements of DBT, ACT, emotion-focused therapy and behaviour therapy. Participants were randomly assigned to this group plus outpatient TAU (N = 12), or outpatient TAU only (N = 10). There were significant between-group differences in favour of the target intervention on most of the study outcome variables including deliberate self-harm, BPD symptomology, depression, anxiety, and EA, and significant pre to post changes in most outcome variables for the intervention condition but not the control condition.

Although these results are promising, particularly for such a time-limited intervention, they are based on a small number of participants, making generalisation from the results difficult. The sample size was also not large enough to allow for formal investigation of variables that might be mediating the therapeutic gains, although emotion dysregulation and EA did significantly reduce, pre- to post-group. Another limitation of the study was that because the intervention combined at least four treatment approaches, it is impossible to determine which are the active components. While it is appropriate to use small samples in early stage treatment development trials, particularly with high-risk patient groups, this study has not yet been followed up by further, larger-scale published trials.

Finally, one study has tested ACT as a stand-alone treatment with a sample that included PD patients. Clarke, Kingston, James, Bolderston, and Remington (in prep), carried out a small-scale RCT comparing ACT (N = 26) with a CBT-style TAU intervention (N = 19), with a heterogeneous sample of patients with treatment resistant mental health problems. Approximately 50% of the sample had a PD diagnosis in addition to Axis I diagnoses. Self-harming behaviour within the previous six months was an exclusion criterion for the study.

Immediately post-intervention, patients in both conditions showed significant improvements in Axis I and Axis II symptomology, and quality of life, with no between-group differences. However, these improvements tended to be better maintained at 6-month follow-up by those who had received ACT, with many in the CBT-TAU condition relapsing. Attrition rates were also significantly higher for the

CBT-TAU condition. Despite these promising findings, the authors noted that the participants with PD symptomology tended to be the most difficult to treat. The majority of PD participants met the diagnostic criteria for just one PD, and at follow-up, 2/3rds of those in the ACT condition no longer met diagnostic criteria for a PD.

As with the other ACT for PD trials, this study was based on a small sample and was statistically underpowered. Nevertheless, the results show that people with complex, chronic mental health problems, including those with PD diagnoses, can benefit from a time-limited, group-based, ACT intervention.

Although not a test of ACT with PD, Berking et al. (2009) found that EA adversely impacted reduction of depression during DBT treatment of BPD, with those participants showing high levels of EA being left with high levels of depression post-therapy. This finding suggests that ACT, a therapy that specifically targets EA might indeed be an appropriate candidate for a post-DBT intervention for PD patients.

In summary, there is a small body of empirical evidence suggesting that ACT-based interventions can be beneficial for people with PD diagnoses. However, no study to date has tested ACT with a heterogeneous group of PD patients (with poor personality functioning across diagnoses), despite co-morbidity in PD diagnoses being common.

CHAPTER III

Methodological Considerations

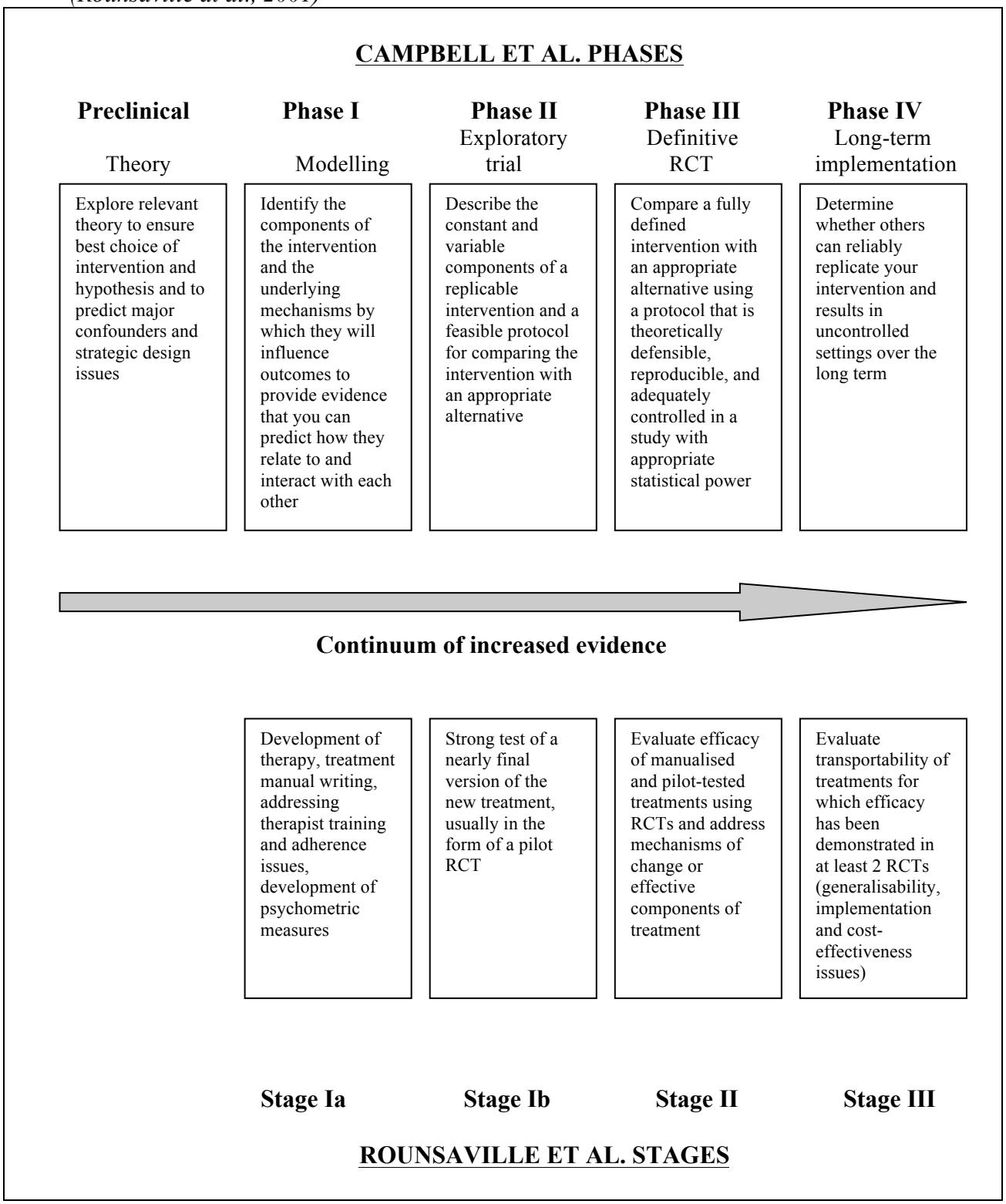
This PhD is concerned with complex psychological problems (PDs), treatment development in relation to these problems, and the particular psychological processes, such as CF, which are thought to underpin these problems and therefore might be expected to be addressed by effective interventions. There is guidance available regarding the process of development and empirical testing of such interventions, guidance that acknowledges the particular methodological challenges presented by this work. For example, the UK Medical Research Council published a framework (Campbell et al., 2000; Medical Research Council [MRC]) guiding methodological decisions through the structure of a four-phase research model, with an additional theoretical, pre-research phase (see Figure 3.1). The authors suggest that the model can be applied to research utilising a range of designs and methodologies, and as such provides a useful adjunct to the large body of clinical research guidance that focuses purely on RCTs (see Nezu & Nezu, 2008, as a recent example).

Alternatively, Rounsaville, Carroll, and Onken (2010) outline a three-stage model that provides a more detailed breakdown of the scientific decisions and activities in the early stages of behavioural treatment development. The stages of this model are mapped onto the Campbell et al. (2000) phases model in Figure 3.1. This chapter will be structured around the Campbell et al. phases, with reference made to Rounsaville et al. where appropriate.

3.1 Preclinical or Theoretical Phase

This pre-research phase involves critically reviewing relevant theories and existing empirical evidence, in order to begin the process of developing an intervention that is thought likely to have a positive impact on the target health problem(s). This stage will not be addresses in detail here, as the previous two chapters of this thesis represent a thorough articulation of this phase.

Figure 3.1
Phases in the Design and Evaluation of Complex Interventions (Campbell et al., 2000) Incorporating Stages in the Development of Behavioural Therapies (Rounsaville at al., 2001)



3.2 Phase I: Modelling

The modelling phase is designed to develop a scientific understanding of the nature and impact of processes relevant to the target problem and the intervention under development, as well as to examine the impact of sub-components of complex interventions (Campbell et al., 2000). It is also at this pre-intervention stage, (Stage Ia in the Rounsaville et al., 2001 model), where appropriate measures to be piloted in the exploratory phase are selected, with new measures being developed as required. Several of the studies in this PhD are designed to serve these various functions, and thus the methodological and statistical issues commonly encountered in modelling research will be examined in detail.

3.2.1 *Measurement*

Selection and/or development of well-designed and methodologically appropriate measures at this pre-clinical phase is extremely important to ensure good quality data and therefore the accuracy of the conclusions drawn (Fernandez-Ballesteros & Botella, 2008). Table 3.1 summarises commonly used types of measures in psychological research. These measures' defining characteristics and their suitability for particular uses, will be outlined below, with particular focus on the types of measures used in this PhD. The methodology for the development of new, self-report questionnaires will also be outlined in detail, as this relates to one of the current studies (see Chapter IV).

3.2.1.1 *Self-Report Measures*

Fernandez-Ballesteros and Botella (2008, p. 95) define *self-report measures* (SRMs) as those that utilise "verbal information about any event reported by a given subject about him- or herself". SRMs are used in the majority of clinical outcome research, and are the only data collection method in approximately 25% of clinical outcome studies (Lambert, 1994). This is despite the widely reported principle that the most likely path to accurate assessment in research, including clinical research, is through the use of multiple assessment methods (Campbell & Fiske, 1959; Eid & Deiner, 2006). The reasons for reliance on SRMs will be discussed in Section 3.2.1.2.

SRMs can be used to collect a wide range of subjective information from research participants, including information about external events (that could be observable by others), as well as internal emotional, cognitive, and physiological events.

Table 3.1

Measurement Methodologies and Data Output

Category of measure	***Type of measure***	***Output data***
Self-report	Psychometric measures such as questionnaires	Quantitative
	Structured interview	Quantitative
	Semi-structured interview	Qualitative or quantitative
	Focus groups	Qualitative
	Diaries/journals	Qualitative
	Self-observation/ self-monitoring	Qualitative or quantitative
	'Think-aloud protocols'	Qualitative
Observation	Behavioural observation	Quantitative or qualitative
	Physiological observation	Quantitative
	Psychological ability tests	Quantitative
	Implicit tests	Quantitative

3.2.1.2 *Psychometric Measures*

Psychometric measures such as questionnaires are the most commonly used measures in clinical research (Fernandez-Ballesteros & Botella, 2008). They are often used to operationalise psychological constructs such as depression or quality of life, and tend to consist of a number of questions or statements, known as *items,* that

address particular symptoms or aspects of the target construct. They can also be used to gather data on opinions and experiences. They are used as *outcome* measures in clinical research, as *process* measures, to examine possible mechanisms of change, and as *screening* measures, in the initial stages of clinical studies.

Self-report psychometric measures are typically self-administered in the absence of the researcher, as this is a relatively economical method of gathering data, (Edwards, 2010). Other advantages to gathering data in this way include the likelihood of reduced social desirability and interviewer influence effects.

However, these advantages may be offset by disadvantages such as more participants failing to complete self-administered psychometrics and thus rendering the data unusable or of questionable quality. Other disadvantages of self-report psychometric measures are features of their design. They tend to consist of closed questions with a small, pre-determined range of possible responses. Clearly, the advantage of this is the standardised nature of the data yielded. However, respondents cannot respond accurately if the measure does not include a response option that matches their experience. Naturally, any measure, including self-report psychometric measures can vary in quality, with poorly designed and tested measures yielding misleading or unusable data. For this reason, there is a lengthy and complex process recommended for the development of new psychometric measures, as outlined below. (e.g., Clark & Watson, 2003).

3.2.1.2.1 Item Development

Nunally and Bernstein (1994) indicate that the first step in developing a psychometric measure of a construct is to use relevant theory to delineate the range of the content of that construct, which then guides the content of the items. Clark and Watson (2003) suggest that the initial item pool should be broader than theoretical understanding of the construct, and include items that ultimately will be excluded for being only tangentially related to the construct.

The specific wording of items and the format of the overall questionnaire are important issues that will impact the performance of the measure. A detailed examination of these issues is beyond the scope of this thesis, but comprehensive guidance is available (e.g., Kline, 1986; Clark & Watson, 2003; Edwards, 2010).

3.2.1.2.2 Psychometric Evaluation

Once a prototype version of a measure has been constructed, it is administered to a suitable sample, and a complex process commences with the aim of examining the performance of the items, the *factor structure* of measure, and its *reliability* and *validity*. Throughout this iterative process, items are excluded from the measure if they are shown to be adversely affecting the psychometric properties of the measure, or are redundant.

The distribution of responses for each item is examined for skew and kurtosis. Items are also dropped if they do not correlate adequately with enough other items or have high enough item-total correlations to warrant further analysis (Field, 2005; Wicksell et al., 2008). At this stage it is also usual to carry out a Kaiser-Meyer-Olkin test of sampling adequacy (KMO; Kaiser, 1970; Kaiser 1974), and Bartlett's Sphericity test to ascertain if the pattern of correlations amongst the items is sufficiently compact and that the items are sufficiently related, to warrant further analysis.

3.2.1.2.3 Factorial Structure

Factor analysis is used to examine the relationships between observed variables (responses to items), and the latent constructs (referred to as *factors*) that it is hoped the measure actually measures (Byrne, 2010). *Exploratory factor analysis (EFA)* is used when this relationship is uncertain, as is the case early in the process of psychometric measure development. Another form of factor analysis, *confirmatory factor analysis (CFA)* will be discussed in Chapter IV.

3.2.1.2.4 EFA

As outlined above, EFA is used to explore the underlying factor structure of a measure. It is also used as a basis for further item elimination. EFA examines whether the relationships between a set of items, that is, common variance, is based in their relationships with underlying factors. There are several available approaches for 'extracting' these factors from the data; the two most commonly used, according to Russell (2002), are *principal components analysis (PCA)* and *principal axis factoring (PAF)*. Based on a review of the evidence, Russell recommends PAF, as it may more accurately reflect population factor loadings than PCA. Furthermore, it is usually the

case that CFA is used in the later stages of scale development, but as PCA is not based on the *common factor model* (Thurstone, 1947), as both PAF and CFA are, it can be problematic attempting to replicate in CFA, facture structures identified using PCA (Brown, 2006).

Commonly, the decision about how many factors should be retained is guided by the Kaiser criterion, which suggests that all factors extracted that have eigenvalues greater or equal to 1 should be retained. However, this tends to lead to over-retention of factors. Instead, the scree test (Cattell, 1966) is used, which involves plotting a graph of eigenvalues against number of extracted factors. The cut off point for the number of factors retained is indicated by a clear change in the slope on the graph.

For ease of interpretation the extracted factors are then rotated. There are various forms of rotation available. Russell (2002) recommends using an oblique (Promax) rotation, which allows for correlation between factors, though in fact the Promax rotation initially involves a Varimax (orthogonal) rotation. This is then followed by a relaxation of the requirement that the factors are uncorrelated. Therefore, the rotation would also indicate if the factors were in fact uncorrelated (Fabrigar et al., 1999).

3.2.1.2.5 Reliability

Field (2005, p.743) defines reliability as "the ability of a measure to produce consistent results when the same entities are measured under the same conditions". The most common measure of scale reliability is *Cronbach's alpha* (α), which is a measure of the internal consistency of the scale. For example, someone who is depressed should get a high score on a psychometric measure of depression. The measure is said to have good internal consistency if that person also scores highly on any randomly selected items from the scale. Caution is needed when interpreting Cronbach's α, as it varies in relation to the number of items in the scale (Cortina, 1993).

The other commonly examined form of scale reliability is *test-retest* reliability, which indicates how stable the measure is over time. The magnitude of the relationship between mean respondents' scores from two time points is usually expressed in terms of Pearson's correlation coefficient (r), with a larger value of r indicating greater reliability.

3.2.1.2.6 Validity

A measure can be demonstrated to be reliable without it being clear that it is measuring the construct it was designed to measure (the *validity* of the measure). Indeed, reliability is necessary but not sufficient for validity in psychometric measures. Therefore such measures should be evaluated against standards of various forms of validity. Some aspects of validity should be addressed at the item generation stage of questionnaire construction. *Face validity,* for example, is an indication that the content of items look like they represent the construct underlying the measure, whereas *content validity,* is the extent to which the domain of content of the construct is sampled through the items. There are no objective ways of measuring or testing face or content validity, but following published guidance on item development and consulting the relevant literature and experts in the field with regards to item content, may increase these forms of validity.

Other forms of validity, such as *predictive, criterion, concurrent, convergent, construct,* and *discriminant validity,* are examined once the measure has been constructed, and can all be objectively measured and tested using statistical means. Nunally and Bernstein (1994) argue that predictive, criterion and concurrent validity are all based on the same logic and procedures and that it therefore does not make sense to treat them as separate entities. Essentially, these forms of validity measure the ability of the scale under development to predict a criterion external to the measure, either using a current criterion or by predicting score or performance on a criterion at some point in the future.

Construct validity (also known as *factorial validity*), is a measure of the relationship between the manifest variables (scale items) and the underlying construct or factors. There are a number of ways of examining construct validity. Convergent validity, for example, is established by testing the strength of relationships (using Pearson's correlation coefficient), between scores on the scale under development, and scores on measures of related constructs. For example, it would be predicted that scores on a measure of CF would significantly correlate with scores on a measure of mindfulness, as these two constructs are related, according to ACT theory, though sufficiently differentiated so as not to lead to a total correlation (1.0) between scores on two such measures. On the other hand, discriminant validity involves demonstrating that there is no significant relationship between scores on measures of constructs that should not, according to theory, be related. It is clear that ideally,

correlations in tests of convergent validity will be high, and those in tests of discriminant validity will be low, though there is no consensus about what constitutes high and low enough (Trochim, 2006). At the very least, discriminant correlations should always be lower than convergent correlations.

Another common approach to examining construct validity is the use of CFA. A model of the relationships between the observed variables (items) and latent variables (the target construct) is developed, based both on theoretical predictions and research findings, including perhaps the outcome of EFA. CFA, in the form of SEM is used to test how well the model fits the data (see Section 3.2.2.3 for an outline of SEM).

3.2.1.2.7 Psychometric Measures in Clinical Research

There are some specific issues to be taken into account when developing measures for use in clinical settings. It is important to administer the measure to the relevant clinical populations during the development process, as item performance, overall measure performance, and factor structure can all vary from sample to sample (Marks, 2004). The measure will need to be stable over time but at the same time, sensitive to change, if being used in an intervention study. Finally, the measure should be able to discriminate between different clinical populations.

In summary, SRMs are the most commonly used measures in clinical research, and have several advantages including ease and cost of administration, as well as the standardised nature of the data they yield, allowing comparison with the findings of other research. However, they can be costly and time-consuming to develop, and the limited response option they offer could lead respondents to ignore item or respond inaccurately (Wilkinson, Joffe, & Yardley, 2004). Other self-report approaches such as interviews might address some of these concerns, and could be considered as alternatives to questionnaires.

3.2.1.3 *Interviews*

Interviews vary between being completely structured and consisting of closed questions, through being loosely (semi) structured and based on open questions that are adapted and extended depending on the material that is arising in the particular interview, to, at the other extreme, interviews that are virtually unstructured. The latter are more common in clinical treatment settings than in clinical research

(Widiger, 2008). It has been argued that structured interviews are essentially "questionnaires that are administered verbally" (Fylan, 2005, p. 65), but with the advantage of the possibility of a rapport being formed between interviewer and interviewee that might aide eliciting of more complete data (Wilkinson et al., 2004). Usually, structured interviews are designed to yield quantitative data, for example structured clinical interviews designed to lead to a categorical clinical diagnosis, such as the Diagnostic Interview Schedule (DIS; Robins, Helzer, Croughan, & Ratcliff, 1981).

Completely structured interview schedules are fairly uncommon, it being more typical for a clinical diagnostic interview to have a fair amount of structure, in terms of the order and wording of questions, but also allow the interviewer some flexibility to ask follow-up and clarifying questions. One example of this type of semi-structured interview offering an advantage over questionnaires is in the field of PD diagnosis. There are questionnaires available, such as the Millon Multiaxial Clinical Inventory III (MMCI-III; Millon, 1994) that are designed to indicate whether an individual has a particular PD diagnosis. However, these types of questionnaires make it difficult to distinguish between occasional experience of symptoms, and long-term difficulties (Perry, 1993) and can lead to 'false positive' diagnoses.

In an attempt to address this issue, other PD diagnostic systems consist of a screening questionnaire, followed by an interview. The interviews use a structured format and proscribed wording of questions, but with some flexibility, to elicit more detailed information on which to base the final diagnosis. The Structured Clinical Interview for DSM-IV Axis II (SCID-II; Spitzer et al., 1996) is an example of this type of diagnostic tool. These kinds of diagnostic interviews are used both as screening tools in clinical research and as outcome measures.

Other kinds of semi-structured interviews are used in clinically focused research. They are often much less structured than diagnostic interviews, and are designed to elicit qualitative data relating to the interviewee's subjective experiences (Boyatzis, 1998). The data they elicit cannot be used to test hypotheses, but one use of such interviews is to form hypotheses that can be tested with further research. This thesis does not involve the use of qualitative analysis methodologies to analyse data from such interviews, and therefore the relevant methodological and design issues will not be reviewed (see Marks & Yardley, 2004 or Forrester, 2010 for an overview).

3.2.1.4 *Other Self-report Measures*

As was outlined in Table 3.2, there are several other self-report measure methodologies used in psychological research. These include the use of *focus groups*, to elicit qualitative data from a group of individuals who have an experience in common (Wilkinson, 2005) and the use of diaries and journals to elicit similar qualitative material on an individual basis (Ferguson, 2005). There are also much less commonly used self-report methodologies such as *self-observation* and *thinking-aloud protocols* (see Fernandez-Ballesteros & Botella, 2008). However, they are not relevant to this thesis, and therefore will not be reviewed.

3.2.1.5 *Observation*

The other main category of measurement methodologies covers those that involve observation of the research participant, rather than their self-report of subjective experience. Due to space constraints, just one observation methodology will be reviewed here; behavioural observation.

3.2.1.5.1 Behavioural Observation

Behavioural observation can be carried out in a structured and relatively objective way, resulting in quantitative data, such as the frequency of a target behaviour. It can also be used in *participant observation* studies, where the influence of the researcher's own perspective and experience is acknowledged (Ballinger, Yardley, & Payne, 2004). Behavioural observation can be carried out in the context of a number of different research designs, including naturalistic designs, for example where naturally occurring interactions between nurses and patients are video recorded or directly observed (e.g. Manias, Botti, & Bucknall, 2002).

Behavioural observation can also be carried out in more contrived situations, where the participant is observed whilst carrying out a standardised activity. In a clinical assessment setting, the aim of this kind of observation is to "derive valid inferences about how the patient will behave in a current or future natural environment" (Haynes, 2003, p. 236). This type of behavioural observation is also frequently used in psychological experiments, for example gathering participant gaze-direction and reaction-time data in a computer-based attentional bias task (e.g. Mogg, Bradley, Field, & De Houwer, 2003). This type of experimental design, often used in the modelling phase of clinical research to examine the role of psychological

processes under controlled conditions, typically uses both observation and self-report data gathering methodologies. An account of the experimental designs and statistical tests used in this type of research can be found in Section 3.2.2.

As with all measurement methodologies, appropriateness and sensitivity to target behaviours, as well as reliability and validity are important issues. Haynes (2003) notes that in the case of clinical assessment, because analogue behavioural observation instruments are usually idiographic, they can have high content validity, are likely to be sensitive to measure the behaviour in question, and acceptable to the individual patient. However, too little attention tends to be paid to other types of validity, and reliability, and of course it is difficult to compare data yielded by different idiographic measures.

Standardised, particularly computer-based observational methodologies have the advantage of collecting data in an objective and accurate manner, and in such a way that it can be compared across participants. However, as is the case with any situation when collecting data in a controlled and contrived manner, the extent to which the data can be seen to represent real life must be considered. Overall, observational data collection can have advantages over self-report methodologies, in that the data is not influenced by the opinions or memory of the participant. However, observation may be time-consuming compared to self-report measurement, as can be the development of such measures.

3.2.1.6 *Conclusions*

This section has reviewed the main measurement methodologies relevant to this thesis. Given that there are so many ways of gathering data, suitability to purpose and quality of the specific measures under consideration should be used to make measurement decisions. In practice, it is desirable to use a number of different measurement methodologies to examine a process or evaluate an intervention, in order to increase confidence in the validity of the process or the intervention (Widiger, 2008; Campbell & Fiske, 1953).

3.2.2 *Model Building and Testing*

In this modelling phase, research studies are often designed to understand the relationships between two or more variables, for example, to establish that a

theoretically consistent relationship exists between a psychological process, such as psychological flexibility, and a psychological problem such as depression. Hayes et al., (2006) review this type of data in support of the ACT model of psychopathology. At the most basic level, this involves establishing the existence of a correlational relationship between the variables, although it is common for researchers to seek a somewhat more detailed understanding of the relationship, by testing whether one variable predicts or is a risk factor for another, (Nock, Janis, & Wedig, 2008).This is commonly tested using regression analysis, where (in simple linear regression), a value of the outcome variable is predicted from a predictor variable, by applying a model to the data that offers the most accurate predictions. This basic approach is extended, in multiple regression, to examine the relative impact of several predictor variables on the outcome variable.

3.2.2.1 *Cross-Sectional Designs*

Both correlational and regression-based studies tend to be associated with *cross-sectional* research designs, involving gathering data on the variables of interest on one occasion only, with all data being collected at the same time. Advantages to this kind of research design include the relative ease, speed and low cost of conducting this type of study. Such studies can be used to generate hypotheses about the relationship between variables, which can be tested more rigorously in subsequent research. The most important limitation of this design is that it is impossible to infer causality in any relationships identified between variables, as such designs do not involve randomisation or collecting data at multiple time points.

3.2.2.2 *Mediation and Moderation Analyses*

One specific use of regression analysis is to test the role of possible mediator and moderator variables. Kazdin (2007; p. 3) defines a mediator as “an intervening variable that may account (statistically) for the relationship between the independent and dependent variable”, and a moderator as “a characteristic that influences the direction or magnitude of the relationship between and independent and dependent variable”. Studies that examine the role of possible mediators or moderators are commonly used in one of two ways. They can be used at the pre-intervention phase to build models of the relationship between relevant variables. For example, Rosenthal, Polusny, and Follette (2006) examined the relationship between perceived criticism

in family of origin (the independent variable; IV) and psychological distress in adulthood (the dependant variable; DV). They concluded that EA fully mediated this relationship.

Mediation analysis can also be used to increase understanding of mechanisms of change in intervention studies. For example, Gaudiano and Herbert (2005) conducted an RCT comparing TAU with TAU plus ACT, for inpatients with hallucinations or delusions. Mediational analysis indicated that believability of hallucinations fully mediated the relationship between frequency of hallucinations and distress. Examination of potential moderators of outcome in RCTs can help identify which patients might benefit most from the intervention being tested, and under what circumstances (Kraemer, Wilson, Fairburn, and Agras, 2002).

Mediation Baron and Kenny (1986) describe the following method, known as the *causal steps approach,* (see Figure 3.2) for testing whether the relationship between an *initial variable*, X, and an *outcome variable*, Y, is mediated by an *intervening variable*, M. The mediational model is specified based on theoretical predictions of the relationships between the variables, and any previous, relevant research findings. The paths between the various variables (a, b, c, c') are then estimated using multiple regression. Other methods of estimating paths in mediation analyses, such as *Structural Equation Modelling*, will be reviewed later.

The Baron and Kenny method involves four steps:

Step 1. Show that X significantly predicts Y by estimating and testing path c, using regression analysis (path c in Figure 3.2: 1.).

Step 2. Show that X significantly predicts the mediating variable, M, by estimating and testing path a (Figure 3.2: 2.), by effectively treating M as an outcome variable.

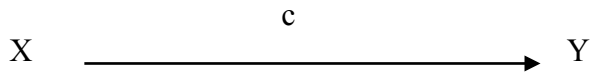
Step 3. Show that the mediator, M significantly predicts Y, by estimating and testing path b (Figure 3.2: 2.). It is not enough to demonstrate a correlation between M and Y, as they might be correlated as a result of both being caused by X. Therefore, X must be controlled when testing the impact of M on Y.

Step 4. Complete mediation by M of the relationship between X and Y is seen to have been established if path c' (Figure 3.2: 2.), drops to zero. That is, if X no longer affects Y when both are entered into a multiple regression model and M is controlled. If the impact of X on Y reduces when M is entered in the model, but is different from zero, M is seen to be partially mediating the relationship between X and Y.

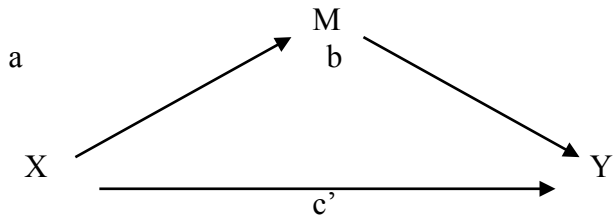
There have been criticisms of the Baron and Kenny causal steps approach (Bollen & Stine, 1990; MacKinnon, Lockwood, Hoffman, West, & Sheets, 2002; MacKinnon & Fairchild, 2009). MacKinnon et al. (2002) have demonstrated that this approach lacks statistical power to detect mediated effects.

Figure 3.2 *Mediational Model Based on Baron and Kenny (1986)*

1.



2.



Bollen and Stein (1990) have argued that the approach is only suitable for variable data that is normally distributed, despite the fact that this is often not the case for mediated effects. The causal steps approach also requires a significant relationship between the predictor (X) and outcome (Y) variables, despite several authors arguing that it is possible for mediation to exist without this relationship being significant (Preacher & Hayes, 2004; MacKinnon & Fairchild, 2009). These criticisms imply that the Baron and Kenny approach is a particularly conservative method, and that there might be situations where mediation effects are occurring but remain undetected. For these reasons, newer approaches to mediation have been developed (see Section 3.2.2.3).

To test the statistical significance of the mediated effect of M, this effect is divided by its standard error, with the result being compared to the critical value, for example $p = .05$. This is the basis of the *Sobel test* (Sobel, 1982), a commonly used, though conservative (MacKinnon, Warsi, & Dwyer, 1995) test of significance in mediational analyses. However, given the likelihood of nonparametric distributions outlined above, and also that the Sobel test works best with large samples,

bootstrapping methodology is increasingly utilised as an alternative means of testing the indirect effect (Preacher & Hayes, 2004). Bootstrapping is a computer-intensive resampling procedure, which is not dependant on the normal distribution of data. In bootstrapping, the sample under analysis is assumed to represent the population. Many (commonly 1000 or more) sub-samples the same size as the original sample are drawn from the original data, with replacement, and used to derive the sampling distribution of the indirect effect. The multiple estimates of the mediation effect are sorted from high to low, and this distribution is used to derive required confidence levels.

There are concerns about the interpretation of mediational analyses, particularly in relation to intervention trials (Kraemer et al, 2002; Kazdin, 2007; Nock et al., 2008). Nock et al. (2008, p. 212) summarise the main concern thus: "just as correlation does not equal causation, mediation does not equal mechanism". By this they mean that although a mediator variable is often referred to as a *mechanism of change*, there is considerable difference between the two. As stated before, a mediator is "an intervening variable that may account (statistically) for the relationship between the independent and dependent variable", (Kazdin, p. 2007; p. 3), whereas, a mechanism is "the basis for this effect, i.e., the processes or events that are responsible for the change; the reason why change occurred or how change came about" (Kazdin, 2007). Demonstrating a statistically significant mediation effect is an important first step in identifying a mechanism of change, but it is not sufficient. Kazdin (2007) outlines seven criteria for establishing a mechanism:

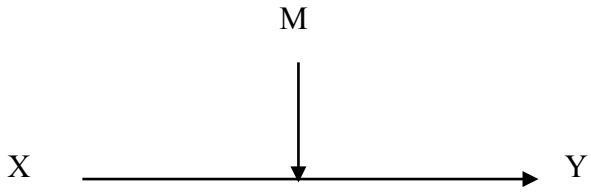
1. Strong association between the relevant variables.
2. Specificity of the impact of the mediating variable on the causal relationship (for example between a psychosocial intervention and anxiety), rather than many variables mediating the relationship.
3. Consistency, in terms of the mediated effect being demonstrated in several studies, with different samples.
4. Experimental manipulation, allowing for greater confidence in the assumption of causality between the initial and the outcome variable.
5. Establishing a timeline, such that initial and mediator variables precede outcomes.
6. Demonstrating the existence of a gradient, such that higher levels (stronger dose or greater activation) of the mediator are associated with greater response in the outcome variable.

7. Plausibility and coherence of the explanation of the mediation effect, based on relevant theory, and other empirical evidence.

These are recent developments in methodological thinking about mechanisms of change, and it is understandable that the majority of psychotherapy research studies (including ACT research), testing mediation, do not yet adhere to these guidelines.

Moderation Where there is a linear, causal relationship between an initial variable X and an outcome variable Y (see Figure 3.3), a third variable, M, can be seen as a moderator of that relationship if it alters the strength of the relationship. For example, a psychosocial intervention (X) may increase quality of life (Y) in patients with chronic pain, but this increase in quality of life might be greater for women than men. Thus, gender (M) moderates the effect of the intervention on quality of life.

Figure 3.3 *Moderation model based on Baron and Kenny, 1986)*



Several key conditions have been identified for moderation effects to be detected and accurately interpreted (Judd & Kenny, 2010), including the need for M to be measured prior to X being measured, particularly if M is a variable that can change. This is less of an issue when M is a variable, such as race or blood type, which does not change. Another important issue is the need to have a robust justification for the direction of the causal (X -> Y) relationship that has been specified, particularly if the initial variable, X, is not manipulated. Multiple regression is used to detect moderator effects by testing the impact of the interaction between X and M on Y.

3.2.2.3 Structural Equation Modelling

As indicated in Section 3.2.2.2, mediational analysis can be carried out using a number of different methods. In recent years, structural equation modelling (SEM) has increasingly been used as an alternative to the multiple regression-based method outlined above. SEM is a general statistical methodology that takes a confirmatory (rather than an exploratory or descriptive) approach to the analysis of a structural model (Byrne, 2010). It has several advantages over more traditional multivariate methodologies, including the capacity to model and test complex, multiple, multivariate relationships simultaneously, as well as the capacity in SEM to assess and take into account measurement errors.

SEM is used to model and test the relationships between both directly observable *manifest* variables, and *latent* variables, that is, variables that cannot be observed or measured directly. In SEM, latent variables associated with IVs are known in as *exogenous* variables (influenced by factors outside of the specified model), whilst those associated with DVs are referred to as *endogenous* variables, and are therefore influenced by the exogenous variables, although not necessarily directly. The hypothesised relationships between variables are specified as a model, and visually represented in a *path diagram*. SEM is used to examine both measurement models, which outline the relationships between manifest and latent variables, and structural models, which examine the relationships just between latent variables.

Kline (2005) outlines six general, iterative steps in SEM:

1. Specify the model. The hypothesised relationship between the various variables is made explicit, usually in the form of a path diagram.
2. Identify the model, by ensuring that it is theoretically possible to calculate all parameter estimates for the hypothesised model.
3. Select appropriate measures of the variables.
4. Estimate the model. This includes calculating the parameter estimates for the model, testing goodness-of-fit of the model, and consider alternative models that could also account for the data.
5. Re-specify the model, if necessary, and go through the analysis steps again.
6. Report the model and analysis thoroughly in any published papers.

SEM is a powerful and flexible analytic tool, used for a range of purposes in addition to its application to mediational analysis, including CFA (see Chapter IV).

Despite the strengths of SEM, as with all methodologies, it has limitations. For example, it requires relatively large sample sizes, with many published studies being based on inadequately sized samples (Westland, 2010). It is also common in published studies for authors to neglect to consider and examine the fit of alternative models. It is also possible, of course, that important variables have been omitted from the model altogether, perhaps due to inadequate attention being paid to theoretical understanding of the subject matter.

3.2.2.4 Experimental Designs

The major limitation with any study based on a cross-sectional design is that it is impossible to infer causality from the results. In fact, assuming causality in a relationship between two variables even in a *prospective or longitudinal* study where data relating to the IV is collected prior to collection of data relating to the DV is still problematic, as it is possible that a third, untested variable might cause both the IV and the DV. The solution, in its simplest form, is to compare two situations, one where the hypothesised cause (IV) is present, the other where it is absent, with all other factors that might affect the outcome being controlled (Field & Hole, 2003). In practice, this is achieved through randomly assigning participants to two conditions and systematically manipulating the IV (Field & Davey, 2005).

This experimental research design is the basis of the RCT, seen as the 'gold standard' for evaluating the efficacy of psychosocial treatments (Nezu & Nezu, 2008). RCTs, including strengths and weaknesses of the design, will be reviewed in Section 3.3.2. However, the experimental design is also the basis of one type of model building and testing research; hence its inclusion in this section. Building models purely based on correlational data is common, due to the relative ease of data collection and the naturalistic set-up of such studies, usually lending the research a degree of *ecological validity* (Field & Davey, 2005). Also, there can be ethical issues involved in research where variables are manipulated in such a way that participants experience detrimental effects, at least temporarily. The difficultly as outlined above, is that such models cannot be said to be truly causal models.

Campbell et al. (2000) also recommend the testing of sub-components of such interventions under controlled, experimental conditions in the modelling phase. As an example, cognitive defusion techniques form part of most ACT treatment protocols,

although relatively little is known empirically about the process of fusion, let alone the efficacy of defusion techniques. A small number of authors (e.g. Masuda et al., 2010) have conducted controlled, experimental studies designed to test the impact of stand-alone defusion interventions.

3.2.2.5 Statistical Considerations in Experimental Designs

The randomisation of participants to two or more controlled conditions, as occurs in experimental research designs, implies that the data from the groups of participants will be compared. The simplest form of group comparison is in the case where there are just two groups, with one IV and one DV. It is usual to analyse the data from such designs using a *t*-test, which generally involves calculating the difference between the observed difference between the group means and the difference between population means (if the null hypothesis was true). This figure is then divided by an estimate of the standard error of the difference between the group means (Field, 2005). There are a number of different forms of *t*-test, depending on whether the same or different participants are assigned to each condition.

In reality, it is rare that clinical experimental research is this simple, and it is more common for some form of *analysis of variance* (ANOVA) to be used, to compare data from several groups, perhaps at more than one time point, or with several variables to be considered simultaneously. In theory, a series of *t*-test could be carried out to analyse data from this type of experiment, but the use of multiple analyses in this way would increase the risk of *type I error,* that is rejecting a true null hypothesis (Wilcox, 2008).

ANOVA is the collective term for a broad and flexible class of statistical models designed to partition the variance (variability in a dataset), and thus to assess the contribution of each variable to the variance. ANOVA can be used in relatively simple situations, for example to examine whether several means are equal or not, in relation to the same variable (known as a *factor)*, in the case of a *one-way* ANOVA. However, they are often used in more complex situations where several levels are examined in more than one factor. There are several ANOVA variants, including *analysis of covariance* (ANCOVA), in which it is possible to partial out the variance associated with a covariate variable, and *multivariate analysis of variance* (MANOVA), used when there are more than one dependant variables.

As with all statistical tests, ANOVA is based on a set of assumptions, the main ones being independence of observations, homogeneity of variance of errors, and normal distribution of scores. In clinical research settings, the assumption of independence of observations in particular, is often invalid, as outcome measures tend to be administered to the same participant on several occasions. In such situations, *repeated measures ANOVA* models are used to determine whether changes in variable score are a function of time.

ANOVA tests can indicate that there are group differences in a dataset, but not specifically where those differences lie (Field, 2005). For this reason, both *planned contrasts* and *post hoc tests* are used following ANOVA, to determine more precisely the location of these differences. However, conducting post hoc tests (essentially multiple *t*-tests), raises the risk of Type I errors, and so *Bonferroni correction* is employed, in which the acceptable Type I error rate (α) is reduced in proportion to the number of post hoc tests.

3.3 Phase II: Exploratory Trial

According to Campbell et al. (2000), this is the phase in the development of a complex clinical intervention where a protocol is developed for comparing (using a pilot RCT) the target intervention with a suitable alternative. Realistically, conducting even a small-scale RCT places a considerable burden on researchers, due to the many treatment and research matters that have to be addressed prior to running an RCT, including the development of a treatment manual, therapist training materials and methods, and methods of evaluating treatment quality and adherence (Rounsaville et al., 2001), as well as the need to meet NHS ethical and research and development requirements. Rounsaville et al. therefore sub-divide Stage I of their model with Stage Ia being where much of this preparatory work is done, alongside the prototype treatment being piloted (in an uncontrolled setting), with a small number of patients. Stage Ib is essentially a pilot RCT, embarked upon once a final version of the new treatment has largely been settled. Rounsaville et al. (2001) argue that the kinds of activities carried out at Stages Ia and b are vital to the success of the overall treatment development process. The complexities of treatment development will be outlined in detail in Chapter VII. Suffice it to say, this can be a lengthy phase of the overall research process (Rounsaville et al., 2001).

As an example, Goldstein, Axelson, Birmaher, and Brent (2007) demonstrated the appropriateness of an open trial in the early stages of treatment development in high-risk circumstances. They adapted DBT for use with adolescents with bipolar disorder (BD), and despite evidence indicating the effectiveness of DBT for adults with other disorders, due to risks associated with BD, and the young, vulnerable patient group, the authors chose a non-randomised, pre-post design. Such designs allow clinicians to adapt the intervention in response to patient response, whilst still providing structure and scientific rigour. Pilot, open trials also tend to involve small numbers of patients, minimising risk if aspects of the new intervention appear to have an adverse impact.

The major limitation of this kind of research design is that without randomly assigning participants to more than one condition, it is impossible to be completely confident that any therapeutic change is actually due to the target intervention.

3.3.1 *Statistical Considerations in Small-Scale, Open Trials*

It is common to analyse data from uncontrolled, pre-post studies, in terms of group differences from pre-intervention to post-intervention, using *t*-tests or repeated measures ANOVAs (see Section 3.2.2.5). However, due to the small number of participants, it may be more appropriate to use non-parametric statistical tests (tests that do not rely on parametric assumptions such as the data being normally distributed). These tests tend to be more robust than parametric tests and have superior power relative to sample size. Commonly, the *wilcoxon signed-rank test* is used in place of a *dependent t*-test, in pre-post studies where scores on an outcome measure are collected at two time points from the same participants. *Friedman's ANOVA* is used as a substitute for a repeated-measures ANOVA.

Another approach to the issue of analysing data from small numbers of participants is to test for *reliable and clinically significant change,* pre to post intervention, for individual participants (Jacobson & Truax, 1991). For each outcome measure, a reliable change index is determined by subtracting the individual's pre score from their post score and dividing the result by the standard error of difference between the two scores (Christensen & Mendoza, 1986). If this index is greater than 1.96, this is taken (at the $p < .05$ level) to indicate that reliable change has occurred.

Jacobson and Truax (p. 633) suggest that the least arbitrary way of defining clinically significant improvement is that "the level of functioning subsequent to therapy places the client closer to the mean of the functional population than it does to the mean of the dysfunctional population." This approach involves calculating a cut-off point midway between appropriate clinical sample and normative sample means, and observing whether an individual participant's score on an outcome measure crosses the cut-off point following the intervention.

3.3.2 *RCTs*

RCTs are currently viewed (e.g. Chambless & Hollon, 1998) as the yardstick by which to evaluate psychosocial intervention trials. The predominant strength of this design is that randomised allocation of participants under controlled conditions protects against threat to the *internal validity* of the research (Clark-Carter & Marks, 2004), that is, the confidence with which causal inferences can be made from the research results. Factors that might, without random allocation to conditions and manipulation of the IV, make it difficult to accurately interpret the results of a clinical trial include pre-intervention systematic differences in participant characteristics such as age, and differences in history, including illness and treatment history (Nezu & Nezu, 2008). There are other potential threats to the internal validity of intervention trials, such as differential *attrition* rates between conditions.

Additional benefits of using an RCT design at the exploratory phase of treatment development are that the actual trial methodologies, such as the randomisation process, that will be used later in the research programme are trialled, and accurate *power* calculations can be made (see Section 3.3.3).

As with all research designs, there are potential weaknesses and limitations of RCTs. Whilst protecting against some major threats to internal validity, RCTs raise the risk of threat to *external validity*, that is the extent to which the conclusions drawn from the study can be generalised to other settings, with other patients, and where the intervention is carried out by other clinicians (Nezu & Nezu, 2008). It has been argued (Seligman, 1995) that comparing interventions under highly controlled research conditions, and deciding which intervention is the most effective and appropriate to offer in the 'real world', are different questions. The issue of

translating results from often highly controlled efficacy research to 'real world' clinical settings, is one that understandably concerns both researchers and clinicians alike (Persons & Silberschatz, 2002), and will be addressed in more detail in Section 3.5.

Regardless of the validity of these criticisms of RCTs, and others, such as the ethical issues raised by randomising patients to conditions that are thought to be less effective than others, or requiring patients to wait, perhaps months, before commencing treatment, RCTs remain central to the process of developing and testing psychosocial interventions (Campbell et al., 2000; Rounsaville et al., 2001). RCTs are by far the most common design of published clinical outcome trials, with detailed guidance being available to improve the quality of reporting RCTs, in the form of the Consolidated Standards of Reporting Trials statement (CONSORT; Altman, 1996; Altman et al., 2001; Boutron, Moher, Altman, Schulz, & Ravaud, 2008).

3.3.3 *Statistical Considerations in RCTs*

The types of statistical tests typically used in RCT-based efficacy trials have already been reviewed in Section 3.2.2.5. However, particular issues regarding *statistical power*, *sample size*, *effect size* and participant attrition are often raised in relation to RCTs and so will be addressed here (although these issues can apply to other types of research designs). The power of a statistical test is an indication of the probability that an effect will be detected (when there is actually an effect to detect). Failing to detect such effects is referred to a *type II* error. So, power = 1 - β, where β is the probability of a type II error. The power of a test is greater if the hypothesis being tested is directional, or if the study has a within-participants design rather than a between-groups design, and power increases with increased control in a study, increased sample size, and increased effect size (Clark-Carter & Marks, 2004). Of these factors, sample size is often the one most readily influenced by researchers. Cohen (1962; 1988) recommends that as a minimum, power of .8 is attained, which still leaves a 1 - .8 = .2 chance (20%), that a type II error will occur.

It is important (particularly in clinical outcome research) to be able to determine not only if a relationship between variables, or the difference between groups is statistically significant, but also if the effect that has been detected is

meaningful. This is done by calculating the size of the effect (in a specific sample), in a standardised way. Cohen (1962; 1988) has outlined a range of methods for calculating effect sizes for different statistical tests, as well as indicating what constitutes small, medium, and large effects. It is common practice to report effect sizes for non-significant results in underpowered studies (e.g. Goldstein et al., 2007).

One of the realities of psychological research, including tests of psychosocial interventions, is that there will be a loss of participants (referred to as attrition), while the study is running. Patients drop out of psychotherapy, often early on in treatment (Kazdin, 2003), even without the additional demands and disincentives of being a participant in a research study. Attrition can undermine the whole design of the study, for example when significant numbers of participants assigned to a waiting list control condition drop out. Such attrition can also make it more difficult to interpret research findings. For example, it can prove difficult to determine whether participants dropped out of a control condition because they were disappointed not to have been randomised to the target intervention, or because the control intervention involved features that were actually unpalatable to some patients.

This kind of differential attrition between conditions is particularly problematic when analysing data from RCTs. It raises the risk of violating the assumptions on which ANOVAs are based. Furthermore, if such participants are excluded from the analysis of the data (a strategy known as *per protocol* analysis), this risks the loss of the advantages of randomisation, and thus is a serious threat to internal validity (Hollis & Campbell, 1999). Intention to treat analysis (ITT), is a recommended alternative strategy (Hollis & Campbell, 1999; Ruiz-Canela, Martinez-Gonzalez, & de Irala-Estevez, 2000), where data from all participants, including those who dropped out or did not adhere in some other way to the research protocol are included in analyses.

Missing data is one of the problematic effects of attrition, and several strategies for estimating missing data values are available, though it has been argued (Hollis & Campbell, 1999, p. 673), that "clinical trials usually do not collect sufficient data to allow good estimation, and the only commonly feasible options are using the last observed response (carry forward) or assuming that all missing responses were constant". This strategy, often referred to as *last observation carried forward* (LOCF), involves using the last collected data from a participant at all subsequent data-collection points, and is based on the assumption that this will yield a

conservative estimate of the impact of the intervention. Of course, if the intervention would have in fact resulted in a negative impact on the participant (one possible reason why an individual might decide to drop out early in from a clinical trial), then LOCF might result in the intervention being viewed as more benign than it actually is.

3.4 Phase III: Definitive RCT

Suitably powered RCTs are extremely expensive and time-consuming to conduct, and it is usual for research funding bodies to require considerable evidence of the kinds of development work outlined above before providing funding for a large-scale RCT (Lancaster, Dodds, & Williamson, 2004). However, pilot RCTs tend to be underpowered, and almost certainly do not involve large enough numbers of participants to allow for investigations of potential mechanisms of change based on mediation analysis or SEM. It is therefore important for a *definitive RCT* to be conducted, in which a well-defined and piloted intervention is tested against an appropriate control condition, under conditions that would allow the detection of any effects that are present, and therefore would also allow confidence in the conclusion that the intervention is not effective in relation to some outcome measures, if no significant group differences are detected. Definitive RCTs are used to demonstrate the efficacy of an intervention, meaning that there is empirical evidence that the intervention 'works' under controlled conditions (Gilbert & Irons, 2005). It does not necessarily follow that the intervention, when offered to patients in routine clinical settings and under less controlled conditions, will still be effective, an issue that will be addressed in Section 3.5.

3.5 Phase IV: Long-Term Implementation

According to the Division 12 Task Force (Chambless & Hollon, 1998), for a psychosocial intervention to be considered well established, it must be supported by evidence of efficacy from at least two RCTs, conducted by independent research groups. Without independent replication, even an intervention with exceptional outcome data from a well-designed RCT will still only be considered promising,

based on these guidelines. Therefore, Campbell et al. (2000) see replication as a vital component of the final stage of treatment development.

During this phase, where a promising intervention is hopefully tested by several groups of investigators, it is desirable for the quality of the control condition to improve. Typically, in early RCTs testing new interventions, the control condition is often *treatment-as-usual* (TAU) or a *waiting list control* (Chambless & Hollon, 1998), with the quality of TAU varying considerably. Although TAU and no treatment control conditions are seen as acceptable by the Division 12 Task Force, moving onto test a new psychosocial intervention against established treatments raises confidence in the target intervention. Certainly, quality of control conditions is one of the main criteria on which criticisms of the evidence base for new and developing interventions are based. For example, Øst (2008) concluded that neither DBT nor ACT could be considered *empirically supported therapies* (EST), based on the Division 12 Task Force criteria, citing poor control conditions, amongst other criticisms.

In Section 3.3.2, a criticism of RCT design was outlined, based on the view that such trials have little in common with 'real world' clinical settings and practice. Indeed, it has been argued (Pearson & Silberschatz, 2003) that in order to reduce variability within groups in RCTs, many patients with multiple or complex diagnoses are excluded (perhaps up to 70%, according to Westen, Novotny, & Thompson-Brenner, 2004), resulting in RCT participants having little in common with actual patients.

There are two main approaches to dealing with this concern regarding the external validity of RCTs. The first is to argue that supportive evidence from RCTs is necessary but not sufficient to establish that an intervention will be effective under routine clinical conditions, and that the definitive RCT phase in treatment development should be followed by a final phase where the generalisability of the intervention to less controlled settings, with different practitioners, is tested. Unfortunately, this frequently does not happen (Clarke, 2003), with less controlled, *field effectiveness* trials relatively uncommon in the published literature.

Another approach, (Seligman, 1995; Pearson & Silberschatz, 2003) is to argue that alternative study designs, such as the *consumer report* (Seligman, 1995) and field effectiveness studies should be used *instead* of RCTs to answer questions about the usefulness of interventions in 'real world' conditions. Of course, excluding RCTs

from the programme of research designed to develop a new intervention would almost inevitably result in a reduction in internal validity. This issue illustrates a broader tension in clinical research design; every decision regarding study design and methodology will in all likelihood result in both an intended advantage and some unwished-for cost or disadvantage. In this way, all clinical research is a process of creative compromise, with the researcher making methodological decisions based on the best information available at the time, the phase or stage in the research programme, and the specific aims of the study in question.

3.6 The Present Thesis

Chapter I reviews the current literature and research evidence in relation to the scientific understanding of PDs and their treatment. From this review, it is clear that there is less empirical research concerned with PDs than with other common mental health diagnoses. More specifically, despite a growing evidence base for psychosocial interventions for BPD, there are significant gaps in the literature regarding the treatment of other PD diagnoses, and for people with BPD who are behaviourally stable following DBT, but who continue to experience psychologically difficulties.

Chapter II critically reviews the development of behavioural and cognitive behavioural psychotherapies, including third wave therapies such as ACT. There are a small number of empirical studies that suggest that ACT might be of benefit to people with PD diagnoses, though there is currently no published test of ACT as an intervention for a group with mixed PD diagnoses, nor as a post-DBT intervention. Chapter II also reviews the important role hypothesised for CF in the development of psychopathology across diagnoses, but notes the paucity of empirical research testing this aspect of the ACT model, with no research examining the relationship between CF and PD. The lack of a well-designed, broadly applicable measure of CF limits the ease with which the results of such studies can be compared, and indeed limits CF-focused research in general. Given that so little is known about the relationship between ACT and PDs, the studies that comprise this thesis fall into the modelling and exploratory phases, as described by Campbell et al. (2000) and Rounsaville et al. (2001).

Chapter III, the current chapter, critically reviews methodological and statistical issues relevant to the types of research included in this thesis.

Study 1 outlines the clinical validation of a new self-report measure of CF, based on a mixed mental health sample including PD. Study 2, an ACT theory development and testing study, utilises this measure in a cross-sectional design with a community sample, to test CF as a possible mediator in the relationship between predictors of PD and actual personality functioning. Study 3, focusing on theory testing, is an analogue study designed to yield a measure of behavioural avoidance and to shed light on the relationship between CF and behavioural avoidance in a student sample (more extreme forms of behavioural avoidance being common and problematic in PD patients).

Studies 4 and 5 extend this programme of research into the exploratory, clinical phase, by testing an ACT intervention with DBT graduates with mixed PD diagnoses and Axis I difficulties. Given that this is a novel use of ACT with a relatively high-risk population, a cautious approach to these clinical trials was taken. Study 4 tests an ACT group-based protocol in an open, pre-post trial with a small sample, with data being collected pre-intervention, post intervention, and at 6-months post-intervention. Based on the data and experiences from this initial group, the intervention protocol was modified, and Study 5 tests the modified, 24-week group protocol with a small sample, again in an uncontrolled trial. Taken as a whole, this thesis outlines an integrated research programme designed to develop scientific understanding of a key ACT process (CF), and of PD from an ACT perspective, and to use this understanding to begin to develop a theoretically coherent, ACT-based intervention for people with poor personality functioning.

CHAPTER IV

Study 1. The Clinical Validation of a Self-Report Measure of CF

4.1 Introduction

CF plays a central role in the ACT theory of psychopathology. However, to date, no adequate measure of fusion has been published, hindering empirical investigation of the construct. The primary aim of this study was to continue the validation process of a new, self-report measure of CF, with a clinical sample, in order to make available a good quality measure for use in both clinical and research settings.

4.1.1 *CF*

CF (defined in Chapter II, Section 2.2.1.2) concerns the way in which we relate to our cognitions. Fusion with thoughts involves taking them to be reality, being unable to see them from different perspectives, being psychologically 'entangled' with thoughts, and thoughts dominating awareness, emotion and action. CF is hypothesised to significantly contribute to psychological inflexibility, and thus to be an important determinant of psychopathology. CF is similar to processes such as *decentering, thought-action-fusion,* and *metacognitive awareness*, but there are important differences. Decentering involves the acceptance of thoughts and emotions, as well as self-compassion, in addition to the capacity to 'step back' from thoughts and emotions (Fresco et al., 2007), the latter being the aspect of decentering most related to CF. Thought-action-fusion (Shafran, Thordarson, & Rachman, 1996) specifically refers to beliefs sometimes associated with obsessional-compulsive disorder that thinking that something harmful will happen actually increases the likelihood of it happening. The construct is thus much more specific than cognitive fusion, and does not (unlike cognitive fusion) refer to a broad awareness of the process of thinking. Metacognition is seen as the awareness of the process of thinking (Teasdale et al, 2002), and beliefs about thinking (Cartwright-Hatton &Wells, 1997), with the one published measure of metacognition being a measure of beliefs about worry and intrusive thoughts.

Theoretically, CF is viewed as a context-determined process. However, for the purposes of questionnaire-based measurement, it is treated as a trait-like construct, where individual differences can be determined. The inconsistency of this strategy with the behavioural theory on which ACT is based, has been discussed by the authors of the AAQI (Hayes et al., 2004), in relation to EA, who underline the pragmatic utility of the approach.

4.1.2 *Measurement of CF*

Despite the hypothesised importance of CF to psychological health (and to ACT theory), there has been relatively little investigation of the process (Blackledge, 2007), compared, for example, with the research literature investigating EA. Even the most basic information relating to fusion, such as norms for different populations, is not available. It could be argued that the paucity of fusion-focused research is in large part due to measurement issues.

There is no single, commonly used self-report measure of CF, in the way that the AAQI (Hayes et al., 2004) and AAQII (Bond, et al., 2011) are commonly used as measures of EA and psychological flexibility. Instead, several authors have developed measures of CF for use with specific populations, with the items of those measures referring to specific cognitions that an individual from that population might experience. An example is the Stigmatizing Attitudes-Believability Scale (SAB; Hayes et al., 2004), which was designed to measure fusion with stigmatising attitudes of counsellors who work with substance misuse patients. This type of measure can be adapted for use with another, similar population. For example, Taylor (2010, unpublished manuscript) used the same basic structure for a self-report measure of fusion with stigmatising attitudes of mental health professionals working with patients with PD diagnoses. However, because these questionnaires use specific cognitions or attitudes as items (e.g., “Personality disordered clients are demanding, you can never do enough”), they cannot be used to measure CF in a range of populations and situations.

There are also examples (Healy et al., 2008; Masuda et al., 2010) of researchers using a single item (for example, “Rate the extent to which you found the previous statement believable”), as a measure of CF. However, it has been argued

that single item measures are less reliable than those with multiple items (Nunally, 1970).

Measures such as the SAB can be effective in terms of doing the job they were designed to do; to measure the believability of certain thoughts for a particular group of people, but they were not designed to be general measures of CF and cannot function as such. They are not flexible enough in their design to be used with a range of populations or cognitions, and therefore they are severely limited in how much they can contribute to an understanding of CF.

There is another, equally important limitation of this type of measure. They were designed to measure the believability of thoughts or attitudes, which is taken to be equivalent to fusion. Indeed, all current published measures of CF are actually measures of believability, despite CF being defined and operationalised in a much broader way in the ACT literature (e.g., Hayes et al., 1999; Hayes and Strosahl, 2004). In these key ACT texts, believability of thoughts is seen as just one aspect of fusion. For a measure of CF to have content validity in terms of addressing the process as it is described in the literature, it would also need to address the other aspects of fusion, such as the inability to view thoughts from different perspectives, outlined in Section 4.1.1. There is no published measure of CF that addresses the process more fully and accurately in this way.

Furthermore, it is not unusual for these believability measures to remain psychometrically untested, leaving a question mark about the reliability and validity of such measures.

4.1.3 *Cognitive Fusion Questionnaire*

Prior to the current study, a self-report measure, the Cognitive Fusion Questionnaire (CFQ; Appendix B) was developed and validated with several non-clinical samples (Gillanders et al., submitted)[2]. We, (Gillanders, Bolderston and Bond) designed the CFQ to be a measure of the broad construct of CF, rather than merely measuring believability. It was also designed to address fusion with

[2] The initial development of the CFQ occurred prior to this PhD. Due to space limitations, the methods employed in its development will not be outlined in detail here, but are described in detail in Gillanders et al.

cognitions in general, rather than specific thoughts, so that it can be used as a general measure of CF across a wide range of populations.

The CFQ consists of 13 items; nine fusion and four defusion. We generated an initial pool of 44 items, which was rated by members of the ACT Special Interest Group of the British Association of Behavioural and Cognitive Psychotherapists on how well each item represented the construct. Ratings were based on a 4-point scale ('not at all representative', 'a little representative', 'moderately representative', and 'highly representative'). Only those items with a modal rating of moderately or highly representative were included in the original prototype of the questionnaire. The final 13-item version was arrived at through an iterative process of item performance analysis and EFA. Items were designed to operationalise key aspects of CF and defusion, including getting entangled with thoughts (e.g. 'I tend to get very entangled with my thoughts'), the dominance of thoughts over emotions and action (e.g., 'I get so caught up in my thoughts that I am unable to do the things that I most want to do'), believability of thoughts (e.g., 'Even when I am having upsetting thoughts, I can see that those thoughts may not literally be true'), and the capacity to experience thoughts from different perspectives (e.g., 'I find it easy to view my thoughts from a different perspective').

The CFQ is based on a 7-point Likert-type scale, with respondents being asked to rate how true each item is for them, ranging from 'never true' to 'always true'. Items address both fusion and defusion, for example 'I overanalyse to the point where it's unhelpful' and 'I find it easy to view my thoughts from a different perspective'. The measure is scored so that higher scores indicate greater fusion (the defusion items are reverse-scored). The CFQ has a Flesch-Kincaid reading index of 5.4, suggesting that the average 10-year-old child could make sense of the wording of the items (Kincaid, Fishburne, Rogers & Chissom, 1975).

EFA with non-clinical samples yielded two factors, with the fusion items associated with one factor, and the defusion items with the other. As was noted in Chapter II, Section 2.2.1.2, in the ACT literature CF and defusion are understood to be opposite ends of the same construct (Blackledge, 2007). It was therefore important to ascertain whether these two factors were both substantive, implying that they represented fusion and defusion as different, though related constructs, or if the two factors were a result of a method effect, in this case due to systematic differences in item wording between the fusion and defusion sets of items (Brown, 2006). CFA

indicated that a one-factor solution with method effect specified (stipulating covariance amongst the error terms for the defusion items) provided the best fit to the data, supporting the hypothesised view of fusion and defusion being opposing elements of the same construct. This finding indicated that the CFQ should be scored as a total scale, and not as separate fusion and defusion subscales, although it is possible that the questionnaire could perform differently, and yield an alternative factor structure, with a clinical sample.

The CFQ appears to be a highly reliable measure with non-clinical samples. Internal consistency of the questionnaire, as measured by Cronbach's α, is excellent, ranging from .81 to .89, with a mean α of .85. Test-retest reliability, assessed using Pearson's correlation coefficient, is also very good (r = .79), over a period of one month. Normative and validity data are available for the CFQ from several non-clinical samples, with the current number of respondents totalling more than 1000. With regards to norms, mean total score across non-clinical samples is 42.89 (SD = 11.73), with total score ranging from 13 to 85 (the scale has a possible range of 13 to 91).

As can be seen in Table 4.1, the CFQ relates as would be predicted to measures of relevant variables. For example, it correlates positively with measures of psychopathology and correlates negatively with a measure of mindfulness (Five Facets Mindfulness Questionnaire; FFMQ; Baer et al., 2008), thus providing evidence for convergent and divergent validity. The relationship between the CFQ and the Balanced Inventory of Desirable Responding Impression Management Scale (BIDR; Paulhus, 1991), a measure of social desirability in responding, has also been examined in a community sample (n = 47). The correlation was non-significant, suggesting that score on the CFQ is not significantly influenced by socially desirable responding.

One matter of note is that there is a particularly large correlation between the CFQ and the AAQII (Bond et al., 2011). In fact the AAQII, designed to measure psychological inflexibility, has been administered to four non-clinical samples with the CFQ, resulting in correlations of .58, .69, .83, and .85. Field (2005) suggests that a correlation as high as .8, and perhaps even above .9 between two variables would be needed to indicate the possibility of multicollinearity.

Table 4.1
Correlations Between the CFQ and Measures of Relevant Variables (Pearson's Correlation Coefficient)

	Measure						
	CORE	*HADS anxiety*	*HADS depression*	*WHOQOL*	*DLSS*	*AAQII*	*FFMQ*
CFQ	.59**	.54*	.39*	-.47**	-.43**	.69**	-.61*
N	113	144	144	113	167	167	44

Note. ** correlation significant at $p < .001$ level (2-tailed) * correlation significant at $p < .01$ (2-tailed). CORE = Clinical Outcomes in Routine Evaluation (global distress); HADS = Hospital Anxiety and Depression Scale; WHOQOL = World Health Organisation Quality of Life brief measure; DLSS - Deiner's Life Satisfaction Scale; AAQII = Acceptance and Action Questionnaire Version II; FFMQ = Five Facets Mindfulness Questionnaire

In this case, this would raise the possibility that the two questionnaires are measuring the same construct. The ACT model would suggest that CF and psychological flexibility are strongly related, but nonetheless, they are described separately (see Chapter II for a detailed examination of this issue). Additional data is needed not only to further test the performance of the CFQ in relation to the AAQII, but to provide insight into the nature of the relationship between these aspects of the ACT model.

In summary, based on the data currently available, the CFQ appears to be a reliable and valid brief, self-report measure of CF, with a factor structure consistent with the ACT theoretical view of the construct. However, to date, it has not been administered to a clinical sample. There are therefore no clinical norms available, and it has not been tested in terms of reliability and validity with a clinical sample. Furthermore, the factor structure has not been examined with a clinical sample, and thus the final structure of the questionnaire, in terms of whether it is psychometrically meaningful or not to score separate fusion and defusion subscales, remains unsettled.

4.1.4 *Present Study*

The purpose of the present study was to extend the development work outlined above by gathering a range of psychometric data in relation to the CFQ, with a mental health clinical sample. This is an important step in the development of the

measure, because it was designed to be utilised in both non-clinical and clinical research settings. If it were to perform well psychometrically with clinical samples, the CFQ could be used for a variety of purposes, including testing the hypothesised role of CF in the development and maintenance of psychological disorders, and assessing the impact of defusion and more general ACT clinical interventions.

The study research questions were as follows:

With a mental health clinical sample:

1. What is the factor structure of the CFQ?
2. What is the internal reliability of the CFQ?
3. What is the test-retest reliability of the CFQ?
4. Does the CFQ demonstrate concurrent validity?
5. Does the CFQ demonstrate convergent and divergent validity in relation to a range of clinically relevant self-report measures?
6. What are the mental health norms for the CFQ (including its relationship with demographic variables)?

4.2 Method

4.2.1 *Participants*

An opportunity, outpatient mental health sample (n = 183) was recruited from a number of sources within Dorset Healthcare University NHS Foundation Trust (DHUFT), and two Scottish NHS Trusts[3]. These sources included community mental health teams and primary care mental health services, as well as specialist PD, eating disorder and community recovery services. All participants had been assessed by a qualified mental health professional as having a current mental health problem such as mood disorders, anxiety disorders, eating disorders, and PDs.

The sample was 73.2% female (n = 134), with a mean age of 39.5 (SD = 12.80; range = 18 - 68). In terms of ethnic origin, 86.9% (n = 159) identified themselves as white, 1.1% as Asian, .5% of mixed ethnic origin, and 10.9% did not

[3] The Scottish data were collected by Dr. David Gillanders (approved by the University of Edinburgh and Scottish NHS research ethics committees), and shared in anonymised form.

provide this information. In terms of current mental health treatment status, 78.1% (n = 143) reported currently receiving psychological treatment such as counselling.

A sample size of 84 was determined to be adequate to ensure acceptably narrow confidence intervals for the construct validity aspect of the study. A sample size of at least 130 was viewed as acceptable for the analysis of the factor structure (Nunally, 1978; Russell, 2002).

Participant inclusion and exclusion criteria for the study were as follows:

Inclusion criteria

(i) 16 years old and above

(ii) Currently using the above NHS mental health services on an outpatient basis.

Exclusion criteria

(i) Patient is not well enough to make an informed decision about participation in the study, or any possibility that participation may have an adverse affect on the patient's psychological wellbeing.

(ii) Under the age of 16.

(iii) Current psychotic symptomology that might impair capacity to give informed consent, or to complete the questionnaires accurately.

(iv) Learning disability

(v) Other organic disorder that might impair capacity to give informed consent, or to complete the questionnaires accurately.

(vi) Currently participating in other research.

Data from an international, community sample recruited via the internet for another study in this thesis (see Chapter V) were used to examine the concurrent validity of the CFQ. This sample (n = 160) was 73.75% female (n = 118), with a mean age of 30.00 (SD = 11.54) and age range of 16 to 70. Country of origin was reported as the US by 53.13% (n = 85), and the UK by 28.75% (n = 46) of the sample. The majority (83.75%, n = 134) reported being of white ethnic origin. In answer to the following question: 'Have you ever sought treatment for a psychological problem (for example depression, anxiety, relationship difficulties, substance misuse, eating disorder)?', 92 (57.50%) responded positively, forming a self-declared clinical sub-sample.

4.2.2 *Materials*

4.2.2.1. *Cognitive Fusion Questionnaire* (*CFQ*; Gillanders et al., submitted) For details of the measure and its development, see Section 4.1.3.

4.2.2.2. *Brief Demographics Questionnaire*. Participants were asked to indicate their age, gender, ethnic origin, and whether they were currently receiving psychological treatment (see Appendix C).

4.2.2.3 Construct Validation Questionnaires. The following common and well-validated measures were used to examine the construct (convergent and divergent) validity of the CFQ.

The *Symptom Checklist-90-R* (*SCL-90-R*; Derogatis, 1994). The SCL-90-R is a widely used 90-item self-report measure, designed to evaluate a broad range of psychopathology symptoms. Respondents indicate on a 5-point scale (0 = *not at all* to 4 = *extremely*) how much they were distressed by each symptom over the past 7 days. Example of symptoms included are 'feeling fearful' and 'other people being aware of your private thoughts'.

The SCL90-R measures nine symptom dimensions (Somatisation, Obsessive-Compulsive, Interpersonal Sensitivity, Depression, Anxiety, Hostility, Phobic Anxiety, Paranoid Ideation, and Psychoticism). Three global indices (Global Severity Index (GSI), Positive Symptom Distress Index, and Positive Symptom Total) can be calculated from the raw data. The GSI, indicating overall psychological distress, is the SCL-90-R index most widely reported in the literature, and is the one used in the current study.

The mean GSI is .31 (SD = .31) for a non-psychiatric sample, with a mean of 1.26 (SD = .68) for a psychiatric outpatient sample. The measure has good psychometric properties including internal reliability for the nine scales ranging from α = .77 to 90, and test retest reliability ranging from r = .80 to .90 for a 1-week period.

The *Personality Diagnostic Questionnaire* (*PDQ*; Hyler et al., 1988). The PDQ is a screening tool for PDs, consisting of 99 true/false questions based on DSM-IV diagnostic criteria. Examples of items include 'I get special messages from things happening around me' and 'Spending time with family and friends just doesn't interest me'. The questionnaire can be used either as a stand-alone screening tool (as it is in this study), to give an indication of personality functioning (a score of 50 or

more positive responses indicating the likelihood of pathological personality functioning), or positive responses to the questionnaire items can be followed up by structured interview questions to yield PD diagnoses. For this type of screening questionnaire, the PDQ shows good internal consistency (mean $\alpha = .71$ across several samples), and test-retest reliability (mean $r = .67$).

The PDQ contains six items designed to highlight possible poor quality data in the form of under-reporting and other suspect responses. An example of these items is 'Sometimes I get upset' (an under-reporting response if answered negatively).

The *Acceptance and Action Questionnaire Second Version* (*AAQ*II; Bond et al., 2011). The AAQII is a self-report measure of psychological inflexibility, the central ACT-relevant psychological process. It has items addressing the individual's relationship to thoughts and emotions, and the impact of thoughts and emotions on the individual's ability to live life as they would like. Items include 'Worries get in the way of my success' and 'I'm afraid of my feelings'. Respondents are asked to rate how true each item statement is for them, on a seven-point scale ranging from *Never true* to *Always true.*

This 7-item questionnaire has good internal consistency (mean Cronbach's $\alpha = .84$ across several samples) and test-retest reliability ($r = .81$ for a 3-month interval). Mean total score for a community sample is 18.53 (SD = 7.52), and for the only clinical sample reported (substance misusers in the US), mean total score is 28.34 (SD = 9.92).

The *Beck Depression Inventory*, 2nd edition (*BDI-II;* Beck, Brown, and Steer, 1996). This is a 21-item self-report questionnaire measuring the range of cognitive, biological, emotional and behavioural symptoms of depression as listed in DSM-IV. Each of the items consists of a symptom, followed by a list of four statements increasing in intensity on a scale of 0 to 3. For example the symptom 'sadness' is followed by the following four statements; 'I do not feel sad' (0), 'I feel sad much of the time (1), 'I am sad all of the time' (2), and 'I am so sad or unhappy that I can't stand it' (3). Respondents are asked to indicate which statement best describes how they have been feeling over the past two weeks.

A score of 0 – 13 is no or minimal depression, 14 – 19 indicates mild depression, 20 – 28 indicates moderate depression, and above 28 (to the maximum

score of 63) indicates severe depressive symptomology. Good internal reliability (α = .92 for an outpatient sample) and test retest reliability (*r* = .93) are reported.

The *Kentucky Inventory of Mindfulness Skills* (*KIMS*; Baer, Smith, & Allen, 2004). The KIMS is 39-item self-report measure of mindfulness, based on the way in which mindfulness is conceptualised in DBT (Linehan, 1993). It gives an overall mindfulness score, as well scores for four components of mindfulness (mindful observing, mindful describing, acting in awareness, and accepting without judgement). The KIMS has good psychometric properties, including internal reliability (α ranging from .83 to .91) and test-retest reliability (*r* ranging from .65 to .83 over a 2-week period).

The *Automatic Thoughts Questionnaire* (*ATQ*; Hollon & Kendall, 1980). The ATQ is a process measure widely used in CBT research. Respondents are asked to rate on a scale of 1 (*Not at all*) to 5 (*All the time*) how frequently over the last week they have experienced each of 30 negative automatic thoughts. Items include 'why can't I ever succeed?' and 'I hate myself'. People with a diagnosis of depression tend to score above 90 on the ATQ, with non-depressed people scoring below 60 (Derubeis et al., 1990). The measure has good psychometric properties, including excellent internal reliability (Cronbach's α = .97).

The Severity Indices of Personality Problems (SIPP; Verheul et al., 2008). The SIPP is a self-report measure of adaptive and maladaptive personality functioning. It consists of 118 items that fall into five domains; self-control, identity integration, relational capacities, responsibility, and social concordance, with respondents being asking to indicate the extent to which they agree or disagree with each statement as a description of themselves over the previous three months. Examples of items include 'I know exactly who I am and what I am worth' (an identity integration item), and 'I can work with people on a joint project in spite of personal differences', a social concordance item. The SIPP has very good psychometric properties, with Cronbach's α for the five domains ranging from .88 to .94, with a mean of .91. Test-retest reliability for the five domains, (over a period of two to three weeks), ranges from *r* = .87 to *r* = .93.

4.2.3 *Procedure*

NHS ethical approval (Reference number: 09/H0502/78; Appendix D), DHUFT research and development approval, University of Southampton Psychology ethical and research governance approval were all obtained for this study.

Clinicians from a range of mental health teams were given the study information sheet (Appendix E). Recruiting via clinicians, though potentially introducing bias into the sample was seen as clinically and ethically preferable to contacting mental health patients without consulting their key clinician. Potential participants were informed about the study either directly by their key clinician, or by the researcher. In all cases, if mental health service users expressed interest in participating in the study, they were given or sent the participant information sheet (see Appendix E). If they then decided to participate, written consent was required (see Appendix E). The participant information sheet emphasised that participation was on a voluntary basis and specifically that access to clinical services would not be affected by the decision to participate or not.

The study consisted of two phases. The initial phase, designed to gather data for validation purposes, involved participants completing the whole questionnaire pack. Based on previous studies, it was estimated that the questionnaire packs would take no longer than 40 – 45 minutes to complete. Participants recruited during the later phase of the study, were asked to complete just the CFQ and the brief demographic questionnaire, taking approximately 10 minutes. These additional data were used to carry out CFA.

Participants completed the questionnaires in the privacy of their home. Consenting participants were offered help to complete the questionnaires, if required, from the IPTS research assistant. No participants asked for this help.

All questionnaires were labelled with a unique identification number for each participant. Questionnaires contained no participant identifying information. Each participant received a debrief sheet (Appendix E), containing supervisor and student contact details. Participants were encouraged to make contact in the unlikely event that they experienced any distress or problem as a result of participating in the study.

A subset of participants who indicated that they were not currently receiving psychological treatment were sent the CFQ only, three weeks after they completed it for the first time, in order to determine test-retest reliability.

4.2.4 *Analysis Plan*

Stage One of the analysis focused on examining the CFQ individual items particularly in terms of skew and kurtosis, as well as examining the data for outliers. Stage Two utilised SEM to carry out CFA in order to examine the factor structure of the measure. The fit of factor models to the data was evaluated using the following fit indices: χ^2/df ratio, comparative fit index (CFI) root mean square error of approximation (RMSEA), and Akaike's information criterion (AIC).

Although commonly reported (usually in the form of the χ^2/df ratio,) χ^2 is highly sensitive to sample size and is considered to be unnecessarily conservative (Brown, 2006). As a result, several alternative fit indices have been developed It is common practice to report one of the comparative fit indices available, such as the CFI, and a fit index that takes parsimony into consideration, such as the RMSEA. The CFI and the RMSEA (both used in this study) are the indices generally recommended (Bentler, 1990; Brown, 2006; Byrne, 2010). The AIC is used to compare two or more models when a χ^2 difference test cannot be used, due to models not being nested. To indicate good model fit, the χ^2/df ratio should be 2.0 or less (Bollen, 1989), the CFI should have a value close to .95 (Hu & Bentler, 1999), and an RMSEA value of .05 or less indicates good fit. For the AIC, a smaller value represents a better model fit.

Stage Three assessed the internal reliability, test-retest reliability, construct and concurrent validity of the CFQ, as well as producing clinical norms. Internal reliability was tested using Cronbach's α, test-retest reliability and construct validity using Pearson's correlation coefficients, and concurrent validity using independent *t*-tests. PASW Statistics 18 software was used for all data analysis except CFA, which was conducted using AMOS 17 software.

4.3 Results

4.3.1 *Data Screening*

Prior to analysis the data were examined in a number of ways. Raw data from the CFQ were examined to see if any individual items had several missing responses, indicating that perhaps participants had difficulties understanding or answering that particular item. This was not found to be the case. In fact there were just 7 missing responses on the CFQ for the whole of the dataset, spread over 6 items.

Missing data was then examined more broadly. It was found that six (7.7%) of the participants who had been administered the full set of study questionnaires (n = 78) had more than 10% data missing from one measure, and therefore the data for that individual for that measure was excluded from subsequent analyses. All other instances of missing data were dealt with by replacing with the sample mean, as is usual practice (Tabachnick and Fidell, 2001). No participants who had been administered just the CFQ had more than 10% missing data. A total of 10 (12.8%) participants had their PDQ data excluded from analysis based on their responses to the PDQ items designed to indicate under-reporting and other suspect responses.

4.3.2 *Stage One: Item Characteristics, Distribution and Outliers*

Participants responded using the full range of possible responses (1 - 7) for all CFQ items. The mean score and standard deviation for each item are shown in Table 4.2. Frequency distributions for each item of the CFQ were tested for skewness and kurtosis (Table 4.3). West, Finch and Curran (1995), recommended by Byrne (2010), suggest that skew and kurtosis values above 7 indicate problematic, non-normal distributions. Using these guidelines, none of the CFQ items were found to have problematic distributions. Box plots and Mahalanobis distance values were used to identify univariate and multivariate outliers (Byrne, 2010), leading to the exclusion of the data from six participants, resulting in a final sample size of 177.

Table 4.2
Descriptive Data for CFQ Items

Items	Mean	SD
1 My thoughts cause me distress or emotional pain	4.85	1.32
2 I get so caught up in my thoughts that I am unable to do the things that I most want to do	4.54	1.45
3 Even when I am having distressing thoughts, I know that they may become less important in the future	4.16	1.40
4 I over-analyse situations to the point where it's unhelpful for me	5.05	1.57
5 I struggle with my thoughts	5.03	1.46
6 Even when I'm having upsetting thoughts, I can see that those thoughts may not literally be true	3.97	1.43
7 I get upset with myself for having certain thoughts	4.72	1.52
8 I need to control the thoughts that come into my head	4.85	1.64
9 I find it easy to view my thoughts from a different perspective	3.34	1.51
10 I tend to get very entangled in my thoughts	4.97	1.50
11 I tend to react very strongly to my thoughts	4.79	1.47
12 It's possible for me to have negative thoughts about myself and still know that I am an OK person	3.89	1.68
13 It's such a struggle to let go of upsetting thoughts even when I know that letting go would be helpful	5.07	1.52

Table 4.3
Skew and Kurtosis Data for CFQ Items (n = 177)

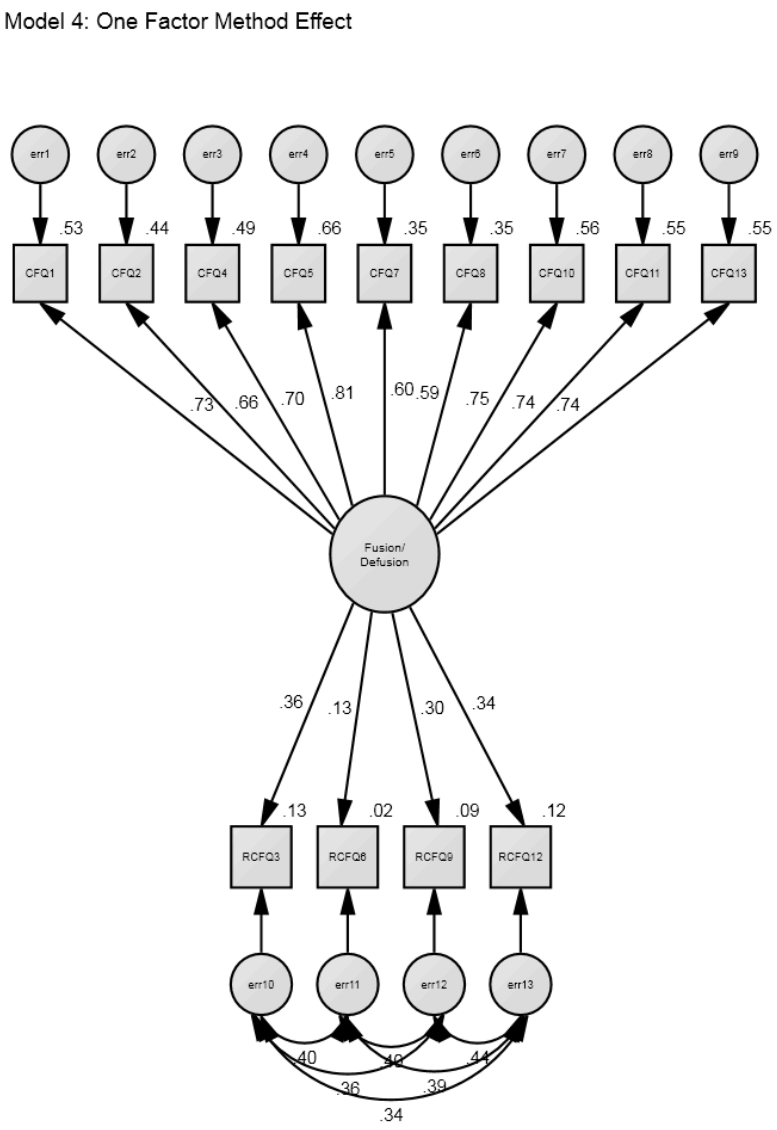
Item no.	Skew	Skew critical ratio	Kurtosis	Kurtosis critical ratio
1	-.26	-1.39	-.06	-.16
2	-.15	-.83	-.09	-.24
3	-.22	1.20	-.03	-.09
4	-.68	-3.70	-.03	-.09
5	-.43	-2.33	-.07	-.19
6	-.08	-.41	-.24	-.66
7	-.48	-2.59	-.07	-.20
8	-.47	-2.55	-.30	-.81
9	-.35	-1.90	-.19	-.52
10	-.68	-3.71	.16	.43
11	-.29	-1.56	-.41	-1.12
12	-.07	-.40	-.63	-1.72
13	-.58	-3.14	-.25	-.68

4.3.3 *Stage Two: CFA*

Based on the results of EFA and CFA with several non-clinical samples, it was predicted that a single factor model with method effect specification would best fit the data. As can be seen in Figure 4.1, the error terms for the four defusion items were covaried, to indicate that the reason why these four items appear to form a separate subscale is due to them all being influenced by a method effect (in this case, systematic item wording differences between fusion and defusion items), rather than them relating to a substantive second factor. Good CFA practice suggests testing a number of theoretically feasible factor solutions (Brown, 2006). Therefore, four models were tested using SEM. Model 1 was a single-factor solution, Model 2, a two unrelated-factors solution, and Model 3, a two-factor solution with related factors. Model 4 was the single-factor model with method effect specification.

From Table 4.4 it can be seen that Models 3 and 4 both fit the data well, with χ^2/ *df* ratios below 2.0, CFI above .95, and RMSEA close to .05. The fit indices for Models 1 and 2 suggest that these two models do not fit the data as closely. A χ^2 difference test cannot be used to assess whether Model 3 or 4 fits the data best, as they are not nested models. However, Model 4, the hypothesised model, provides the best fit on all indices, including the AIC, which is specifically used to compare models. Given, additionally, that this model is more parsimonious and is also theory-consistent, it is the model that should be retained.

Figure 4.1
Model Specification of Predicted Factor Structure with Standardised Regression Weights



Note. All paths significant at $p < .01$. Large circle represents latent variable, squares represent manifest indicators, err denotes error terms. Double-ended arrows indicate covariance between error terms.

Table 4.4
Fit Indices for CFA Comparing Four Factor Models for the CFQ

	Fit index					
Model	χ^2	*df*	χ^2 /*df*	*CFI*	*RMSEA*	*AIC*
1. One factor	201.63	65	3.12	.84	.11	279.63
2. Two unrelated factors	114.50	65	1.762	.94	.10	192.50
3. Two related factors[1]	99.15	64	1.55	.96	.06	179.15
4. One factor, method effect specified	86.88	59	1.47	.97	.05	176.88

Note. χ^2 = minimum fit function chi-squared; *df* = degrees of freedom; *CFI* = comparative fit index; *RMSEA* = root mean square error of approximation; AIC = Akaike's Information Criterion; [1] Pearson's correlation between the two related factors = -.34, p < .001

4.3.4 *Stage Three: Reliability*

Internal consistency, as measured by Cronbach's α coefficient was examined for the CFQ total scale, the fusion subscale and the defusion subscale. The results can be seen in Table 4.5, with equivalent statistics from community samples, for comparison. The total scale and fusion subscale was shown to have excellent internal consistency, while the defusion subscale was shown to have acceptable internal consistency, particularly when the fact that it consists of just four items is taken into account. All α values were above the commonly used benchmark of .70 (Tabachnick & Fidell, 2001).

Test-retest reliability was assessed by examining the relationship between CFQ total scores from questionnaires completed three weeks apart by a subsample of participants (N = 19). A significant correlation was found between the two scores, using Pearson's correlation coefficient (r = .84, p < .001).

Table 4.5
*Cronbach's α Coefficient for the CFQ, for Clinical and Non-Clinical Samples**

	Cronbach's α		
CFQ scale	*Clinical sample n = 177*	*Non-clinical samples Range n = 1072*	*Non-clinical samples Mean n = 1072*
CFQ total scale	.86	.81 - .89	.86
CFQ fusion subscale	.89	.88 - .93	.91
CFQ defusion subscale	.72	.54 - .78	.72

Note. *Non-clinical values quoted from Gillanders et al., (submitted)

4.3.5 *Sample Characteristics*

Table 4.6 shows clinical normative data for the CFQ, presented alongside equivalent data from community samples. As would be expected, this clinical sample had a higher mean total fusion score than the non-clinical sample. Analysis of CFQ score by gender using independent *t*-tests indicated yielded no significant gender differences on the total score ($t(175) = 1.29$, $p = .20$), the fusion subscale, ($t(175) = .73$, $p = .47$), or the defusion subscale, ($t(175) = -1.80$, $p = .07$), although the latter was close to being significant, with men, on average, scoring as more defused than women. The relationship between CFQ score and age was assessed using Pearson's correlation coefficient. No significant relationships were found, for the CFQ total scale ($r = -.09$, $p = .23$), fusion ($r = -.08$, $p = .27$), or defusion ($r = .06$, $p = .45$) subscales. This is in keeping with data from community samples.

Table 4.6
*Normative data for the CFQ, for clinical and non-clinical samples**

CFQ scale	*Clinical sample mean score (SD) n = 177*	*Clinical sample range n = 177*	*Non-clinical samples mean score (SD) n = 1072*	*Non-clinical samples range n = 1072*
CFQ total scale	60.77(11.37)	31 - 86	42.89(11.73)	13 - 85
CFQ fusion subscale	44.46(9.45)	20 - 63	30.49(10.03)	9 - 63
CFQ defusion subscale	15.20(4.30)	4 - 27	12.40(4.22)	3 - 28

Note. *Non-clinical values quoted from Gillanders et al., (submitted)

4.3.6 *Concurrent, Convergent, and Divergent Validity*

It was expected that the CFQ could discriminate between clinical and non-clinical samples, with the latter displaying lower scores (concurrent validity). This was tested using an internet sample from another study that forms part of this thesis (see chapter VI), as that study yielded a self-declared clinical sample (n = 92) and a self-declared non-clinical sample (n = 68)[4], both of which had also completed the same measure of psychopathology, the SIPP (Verheul et al., 2008). This allowed the possibility not only to assess whether the CFQ could distinguish between the self-declared clinical and non-clinical samples, but also to verify the clinical status of these samples by testing for a difference between them in terms of SIPP score.

As predicted, the mean total CFQ score of the clinical sample (51.49, SD = 14.56) was higher than that of the non-clinical sample (41.57, SD = 13.70). Similarly, the mean average SIPP score of the clinical sample (4.89, SD = .81) was lower than that of the non-clinical sample (5.28, SD = .70), indicating poorer personality functioning. There was a significant difference in mean CFQ total scale score between the self-declared clinical and non-clinical samples ($t(158) = 4.37$, $p < .001$)

[4] Participants were asked the following question: 'Have you ever sought treatment for a psychological problem (for example depression, anxiety, relationship difficulties, substance misuse, eating disorder)?'

and between the two samples in terms of SIPP score ($t(157) = -3.16$, $p = .002$). This provides preliminary evidence that the CFQ can distinguish between clinical and non-clinical samples that have been verified in terms of score on a well-validated and reliable indicator of psychopathology.

Further evidence of concurrent validity would come from CFQ scores correlating with measures of psychopathology and distress, as CF is viewed in the ACT model as contributing to these types of difficulty (Hayes et al., 1999). Relationships between the CFQ and such measures were assessed using Pearson's correlation coefficient (Table 4.7). Results were consistent with the hypothesis that score on the CFQ would be positively correlated with measures of psychological difficulties such as depression and personality problems. This was the case both for the CFQ total scale, and for the two possible subscales, which both related to measures of psychopathology in expected directions.

Table 4.7

Correlations Between the CFQ and Measures of Psychopathology (Pearson's Correlation Coefficient)

	Measure		
CFQ scale	*BDI*	*SCL-90*	*PDQ*
CFQ total scale	.71** n = 77	.66** n = 76	.48** n = 62
CFQ fusion subscale	.66** n = 77	.62** n = 76	.47** n = 62
CFQ defusion subscale	-.56** n = 77	-.52** n = 76	-.32* n = 62

Note. ** correlation significant at $p < .001$ level (2-tailed) * correlation significant at $p < .01$
BDI = Beck Depression Inventory; SCL-90 = Symptom Checklist-90-R Global Severity Índex; PDQ = Personality Diagnostic Questionnaire

It was also expected that score on the CFQ would correlate positively with measures of related variables, such as the AAQII, a measure of psychological inflexibility (convergent validity), but would have an inverse relationship with the KIMS (divergent validity). From Table 4.8 it can be seen that this is indeed the case,

again with all forms of the CFQ demonstrating significant correlations with these variables, in the expected directions. The CFQ total scale and fusion subscale both correlate particularly strongly with the AAQII, an issue that has already been raised in relation to the validation of the CFQ with non-clinical samples (Section 4.1.3). Although the *r* values for the clinical sample fall just below those suggested by Field (2005) as possibly indicating multicollinearity, these results and their possible implications regarding the relationship between the CFQ and the AAQII will be addressed in detail in the discussion (Section 4.4.2).

Examination of the correlations between the two potential CFQ subscales and all of the variables in Tables 4.7 and 4.8 shows that the subscales appear to be performing very similarly, although inversely. This provides support for the interpretation that rather than relating to two substantively different factors, the fusion and defusion items are in fact just oppositely worded and scored indicators of a single factor.

Table 4.8
Correlations Between the CFQ and Measures of Relevant Variables (Pearson's Correlation Coefficient)

	Measure		
CFQ scale	*AAQII*	*KIMS*	*ATQ*
CFQ total scale	.78** n = 77	-.67** n = 78	.64** n = 78
CFQ fusion subscale	.74** n = 77	-.61** n = 78	.58** n = 78
CFQ defusion subscale	-.57** n = 77	.54** n = 78	-.54** n = 78

Note. ** correlation significant at $p < .001$ level (2-tailed) * correlation significant at $p < .01$
AAQII = Acceptance and Action Questionnaire, Version II; KIMS = Kentucky Inventory of Mindfulness Skills; ATQ = Automatic Thoughts Questionnaire

4.4 Discussion

CF plays a central role in ACT theory, according to which it makes an important contribution to establishing psychological difficulties. Despite this hypothesised role, there has been little empirical research conducted to examine the nature of CF or its impact on mental health, to a large extent due to the lack of a well-validated, broadly applicable measure. A new measure, the CFQ, has been developed to address these issues, and early data with non-clinical samples has been encouraging (Gillanders, et al., submitted). The aim of this study was to build on this prior CFQ development research by beginning the process of examining the psychometric properties and performance of the measure with a clinical sample. The results of this study, based on an NHS mental heath outpatient sample, provided promising evidence regarding the performance of items, factor structure, reliability, and validity of the measure. The study findings, and their implications including recommendations for further research, will be discussed below, in the order that the research questions were originally stated in Section 4.1.4.

4.4.1 *Study Findings*

4.4.1.1 Factor Structure

A fundamental issue to be addressed in this study was the factor structure of the CFQ. With non-clinical samples, a one-factor model with the specification of a method effect provided the best fit to the data, suggesting that (in accordance with ACT theory), CF as measured by the CFQ is unidimensional. However, questionnaires can yield different factor structures with different types of samples, so it was important to test the factor structure with mental health participants. Two models fit the clinical data well; a two-related-factors model and a one-factor with method effect model. However, the latter provided the best fit to the data, and coupled with the findings regarding performance of the fusion and defusion subscales in relation to other variables, these findings confirmed the single-factor structure observed with non-clinical samples.

This finding is important for a number of reasons. Crucially, it enables the format of the measure to be finalised. It would not make sense to score the fusion and

defusion items as separate subscales, given that they do not appear to represent two substantive factors. As a result, the CFQ will be scored as a total scale only. It could be argued that if the 13 items are indicators of just one factor, given that the defusion items appear to add little to the scale psychometrically-speaking, they could be excluded from the final version of the measure. However, these items address important content of the construct, such as believability of thoughts, which is not addressed by fusion items. Therefore, in the interests of content validity, all 13 items have been retained.

Replicating the factor structure indicates also that the measure is performing consistently across very different samples, including the current, purposely mixed-diagnosis sample. Furthermore, the fact that a single factor structure best fits the data has important implications, in that this is the first attempt to produce a measure of the broad construct of CF, as an alternative to the existing measures that use believability of thoughts as a limited proxy for the construct. The fact that these 13 relatively heterogeneous items load onto just one factor provides initial empirical support for the way in which fusion is conceptualised in ACT theory and interventions.

4.4.1.2 Reliability

The CFQ demonstrated very good internal reliability and test-retest reliability with this clinical sample. On both counts it was found to be as reliable with this clinical sample as it is with community samples.

4.4.1.3 Construct Validity

The CFQ performed well in relation to several forms of construct validity. It was able to distinguish between clinical and non-clinical samples (concurrent validity). It also related in predicted ways to measures of several relevant variables, demonstrating convergent and divergent validity. This included the CFQ strongly relating to well-validated measures of psychological distress and disorders, such as the SCL-90-R (Derogatis, 1994) and the BDI-II (Beck et al., 1996). Although this is cross-sectional, correlational data, from which causality cannot be inferred, these results are consistent with ACT theory regarding the association between CF and psychopathology. Further research, in the form of RCTs and mediation studies, are required to examine a potential causal role for fusion.

It was noted earlier (Sections 4.1.3 and 4.3.6) that with both non-clinical and clinical samples, the CFQ and the AAQII are highly correlated. This is important because it relates to the way in which CF and psychological flexibility are conceptualised and relate to each other, issues about which there is a lack of clarity in the ACT literature. Some authors describe fusion and flexibility as if they are related but differentiated (e.g. Blackledge, 2007), with the implication that it should be possible to measure and investigate them independently. Other authors, for example, Wilson & Dufrene (2008) suggest that there is essentially just one process at work (psychological flexibility), and that terms such as CF and EA represent the different ways in which psychological flexibility can impact human experience. Therefore, attempting to develop a stand-alone measure of CF, and then examining how it relates to the AAQII (designed to measure psychological flexibility), could provide important data to help guide clarification and refinement of this aspect of the model.

With non-clinical samples the correlation between the two measures has ranged from .58, which would seem to support the 'related but separate' view, to .85, which could be seen as an indication that the two measures are in fact measuring the same construct. The correlation between the two measures with the clinical sample was .78; just below the level at which multicollinearity might be suspected (Field, 2005). Unfortunately therefore, this finding does not particularly help clarify the situation regarding ACT processes and how they relate. It could also be argued, that even if a view was taken that both measures are basically measuring the same construct, this does not necessarily imply therefore that there is no difference between CF and psychological flexibility. An alternative explanation could be that one or both of the measures does not accurately reflect the construct it was designed to measure. Specifically, given the questions about the validity of the AAQ and AAQII (see Section 2.2.1.2), it is possible that the high correlation between the CFQ and the AAQII results from problems with the latter measure, and therefore should not be interpreted as indicating that fusion and inflexibility of essentially the same construct.

However, if it was assumed that the CFQ and the AAQII are accurately measuring the same construct, then the CFQ might more precisely be described as a measure of psychological flexibility in relation to cognition. Further research is required to explore these various issues, including an examination of how the two measures each relate to a range of relevant variables, as well as how they both perform in various research and clinical settings. Additionally, the items from both

measures could be entered into EFA together, to see if they load onto different factors.[5]

4.4.1.4 Normative Issues

This study allowed clinical norms to be ascertained for the CFQ for the first time. As expected, mean CFQ score for this mental health sample was higher than the mean score derived from non-clinical data. No relationship was found between score on the CFQ and age, or between CFQ score and gender. Both of these findings fit with those from research with non-clinical samples.

4.4.2 *Methodological Limitations*

The main methodological limitations of this study relate to the sample, in terms of size, composition, and recruitment. As noted in section 4.2.1, a sample size of 84 was determined to be adequate to ensure acceptably narrow confidence intervals for the construct validity aspect of the study. In fact 78 participants were recruited to this part of the study, resulting in it being marginally underpowered. This was a particular issue in relation to the PDQ, which requires some completed questionnaires to be excluded from analysis due to responses on screening items. However, the precise size of correlational relationships is not usually predicted in this type of study, with the direction and significance of the relationship usually seen as most important. Given the significance levels and sizes of correlations between variables found in this study, it seems unlikely that the conclusions drawn from the findings would change substantially, even if some small changes in *r* values resulted from an increased sample size.

Sample composition was an additional issue for this study. Although a strength of the study was that participants were recruited from a broad range of mental health services, increasing the generalisability of the findings, there were other sample composition-related limitations. Men were underrepresented, as were people with non-white ethnicity. The sample also includes few participants above the age of 65. Caution is therefore required when generalising from the results, and future

[5] Since the current research was completed, data from a study investigating distress in prison staff has shown that the CFQ accounts for variance in distress above and beyond that accounted for by the AAQII (Gillanders et al., submitted), thus providing support for the hypothesis that the CFQ and AAQII measure related but distinct constructs.

research will be necessary to evaluate the performance of the CFQ with other, more varied clinical samples.

Finally with regards to sample-related methodological limitations, the study recruitment strategy may have introduced bias in the sample, because clinicians were free to decide not to discuss the study with patients who they deemed unsuitable. This raises the possibility for example, that potential participants with more severe difficulties, more acute difficulties, and particular diagnoses, may have been excluded through the recruitment process. However, a suitable balance has to be reached between requirements of scientific rigour and clinical and ethical issues, and particularly for the first testing of a new questionnaire with a clinical sample, this level of caution was appropriate.

4.4.3. *Future Research*

In addition to the continued validation research indicated in earlier sections of the discussion, rather than using further heterogeneous samples, future research could take a different approach to establishing the applicability of the CFQ to people with a range of diagnoses. Repeated demonstration of the utility of the CFQ with samples of people with more specific diagnoses such as depression and psychosis, would robustly establish the general applicability of the measure. Such an approach, requiring as it would the recruitment of several clinical samples, was beyond the scope of this PhD. The CFQ will also need to be tested in terms of sensitivity to therapeutic change. There is some initial data indicating that scores on an earlier prototype of the CFQ changed significantly following an ACT-based intervention for PTSD (Bastien, Hermann, & Moore, 2010), but this should be verified with this final version of the measure. To this end, the CFQ has been included as a process measure in the pilot treatment development studies that forms part of this PhD (Chapters VII and VIII).

The unique flexibility that the design of the CFQ affords, opens the door for a range of important new research developments. The CFQ could be used to provide data about the nature of CF and its impact on psychological health and ill health, including how different populations, with different experiences, compare with regards to levels of fusion. The measure could also play a role in testing important aspects of the ACT model, including shedding light on the nature of psychological

flexibility, the central ACT process. Finally, it could be used to measure the impact of standalone cognitive defusion techniques, as well as broader ACT treatment protocols, with the possibility, for the first time, of comparing these findings across different samples and diagnoses.

4.4.4 *Summary*

In summary, the CFQ is a brief self-report questionnaire that provides a novel approach to measuring CF, an important ACT-relevant construct. Data from this study show that it performs well with a mental health sample, with very good reliability and validity, whilst having greater content validity than alternative measures. As with all new psychometric measures, an on-going process of validation is required to continue to increase confidence in the measure.

CHAPTER V

Study 2. Modelling the Role of CF in Relation to Personality Functioning

5.1 Introduction

The ACT model indicates that CF plays an important role in the development of psychopathology. Given that it is a transdiagnostic model, the implication in relation to PD is that CF is associated with poor personality functioning in general, rather than a specific PD diagnosis. With the development of the CFQ (as outlined in the previous chapter), it is now possible to model and test these relationships. The current study was designed to test specifically whether CF mediates the relationships between two risk factors of poor personality functioning – negative affectivity (NA) and childhood trauma (CT), and actual personality functioning in adulthood.

5.1.1 *Risk Factors for Poor Personality Functioning*

Both genetic and environmental risk factors for PD have been investigated empirically. Kendler et al. (2008) identified three genetic risk factors for PD, the most substantial of which was labelled “negative emotionality”. This had significant loadings on six PD diagnoses, across all three clusters, and was characterised as reflecting “a broad vulnerability to PD pathology” (p. 1438). In essence, this factor appears to represent a general vulnerability to poor personality functioning in adult life, regardless of specific PD diagnosis. Other researchers have reported similar findings (e.g. Livesley et al., 1998).

Environmental risk factors for PD have also been identified. For example, Johnson et al. (1999) found that traumatic experiences in childhood, such as abuse and neglect, significantly increased the risk of PD in adulthood, a finding that they have replicated and explored in detail (Johnson, Cohen, Chen, Kasen, & Brook, 2006). Again, these kinds of traumatic experiences are implicated across PD diagnoses, suggesting that they increase vulnerability to poor personality functioning in general.

5.1.1.2 *NA and PD*

One way in which negative emotionality (the broad genetic risk factor) has been operationalised is as NA – the temperamental predisposition to experience negative affect (Larsen & Diener, 1987; Bryant, Yarnold, & Grimm, 1996). Several studies have investigated NA as a risk factor for PD-related behaviours. Gratz (2006) found that NA was a risk factor for deliberate self-harm (DSH) (associated with some PD diagnoses), for female students, while Gratz and Roemer (2008) also reported that one aspect of NA, negative affect intensity, was related to DSH. Kingston et al., (2010) found that negative affect intensity was significantly related to engagement in maladaptive behaviours often associated with PD, including DSH and substance misuse.

Lynch, Robins, Morse, and Krause (2001) reported that with a sample of psychiatric outpatients (50% with PD diagnoses), negative affect intensity was significantly correlated with hopelessness and depression (both associated with PD). Lynch, Cheavens, Morse, and Rosenthal (2004) found that with a sample of depressed older adults, NA, suicidal ideation and hopelessness were correlated.

Although these studies involve DVs that are correlates of PD, rather than actual poor personality functioning, taken as a whole, their results suggest that NA is likely to be a risk factor for poor personality functioning.

5.1.1.3 *CT and PD*

CT, in the form of abuse and neglect, has consistently been shown, through both cross-sectional and prospective studies, to be associated with or to predict many psychological disorders (Marx & Sloan, 2002; Widom, DuMont, & Czaja, 2007). This includes a range of PD diagnoses, in both men and women (e.g. Herman, Perry, & van der Kolk, 1989; Raczek, 1992; Krinsley et al., 1992).

5.1.2 *The Role of CF*

Despite the evidence suggesting that both NA and CT are associated with a significant increase in the risk of poor personality functioning in adulthood, not everyone who experiences these risk factors will go on to develop such difficulties. It is therefore important to identify variables that mediate these relationships. Given that

adult patients cannot alter their genetic endowment or childhood experiences, it may be of considerable value to identify intermediary processes that could be addressed through psychosocial interventions.

According to the ACT model, the way that individuals relate to their private experiences including memories of CT and negative emotions – in particular the extent to which they are dominated by these experiences (CF) – will be implicated in psychological suffering (Hayes & Strosahl, 2004). Clinical observation supports the idea that CF might be a mediating variable in relation to PD. Patients with PDs appear to struggle with cultivating a defused relationship with private experiences. Being dominated by their private experiences in this way appears to play a role in maintaining poor functioning, for example by increasing the urge to avoid situations in which fused-with aversive thoughts might be experienced. Clinical practice suggests that as long as PD patients remain fused in this way, it is difficult for them to engage with exposure-based interventions designed to reduce avoidance. Observations of this kind are consistent with the hypothesis that CF leads to EA (Greco et al., 2008).

There is currently no published research testing CF as a mediating variable in the relationship between established risk factors and poor personality functioning, but a small number of studies have investigated related issues, such as the role of fusion as a mediator of other forms of psychological difficulties, and the mediating effects of variables linked with CF. In terms of the role of CF in relation to psychological problems, it has been shown to mediate the impact of ACT on depression (Zettle et al., 2011), and it also appears to be associated with reduction in hospitalisation rates for people with psychosis (Bach & Hayes, 2004).

Several studies have tested EA, as measured by the AAQ (Hayes et al., 2004), as a mediator in relation to the kinds of risk factors discussed above, and in relation to IVs that are correlates of PD. These studies are relevant because of the hypothesised link between CF and EA and the high correlation between the CFQ and the AAQII demonstrated in Chapter IV. For example, Kingston et al. (2010) found that EA mediated the relationships between two risk factors (CT and negative affect intensity) and PD-relevant maladaptive behaviours such as DSH. Similarly, Reddy, Pickett, and Orcutt (2005) reported that EA mediated the relationship between childhood psychological abuse and mental health symptoms in college students. Marx and Sloan (2002) found that in a college sample, EA mediated the relationship between CSA

and psychological distress. The same process has been found to mediate dropout from an inpatient DBT treatment program for BPD (Rusch et al., 2008), and predict depression levels in BPD patients receiving DBT (Berking et al., 2009).

Given the link between EA and fusion, in addition to the clinical observations previously outlined, it seems logical to hypothesise that CF will act as a mediator in relation to personality functioning. A measure designed to address CF in a clear, precise manner should be used to test this, to avoid the kind of confusion that exists in relation to the measurement of EA and psychological flexibility at this time (see Chapter II, Section 2.2.1.3 for details).

5.1.3 *Study Synopsis, Methodological and Design Considerations*

The purpose of the present study was to test CF as a mediating variable in the relationships between NA and personality functioning, and CT and personality functioning. This was done by first conducting a mediational analysis (following Baron & Kenny, 1986) with each risk factor relationship separately. It was predicated that CF would fully mediate both relationships. Following this, mediational models involving both risk factors were tested using SEM.

The design of this study was influenced by the methods available for measuring the variables of interest and the design of relevant published research. There are easily administered, validated self-report measures available for all the study variables. Although there are alternative methods of measuring CT and personality problems (e.g. semi-structured interviews), they are time-consuming, expensive, and do not allow participants to anonymously provide sensitive information. In addition, interview-based assessments of personality problems (e.g., the SCID-II) yield a set of diagnoses rather than a personality functioning continuous variable. For these reasons, the study is based on self-report measures (see Section 5.2.3).

Testing models of the study variables in the manner outlined above, using a cross-sectional design, offers an important first step in assessing the relationships between these variables, and is in keeping with much of the relevant published research. If these models prove useful, other designs that involve randomisation and collecting data prospectively would be helpful to assess causality.

In order to provide a setting in which participants could share personal information privately, as well as to facilitate the recruitment of a large, international sample, the study was based on the internet, being accessed via several public-access research websites.

5.2 Method

5.2.1 *Design*

This study was based on a cross-sectional design, using self-report measures of the variables of interest. The mediational variable tested was CF. The IVs in the study were CT and NA; the DV was personality functioning in adulthood.

5.2.2 *Participants*

An opportunity sample (n = 234) was recruited via public-access research websites (see Appendix F). The only exclusion criteria were that participants had to be at least 18 years old and be able to understand English. As can be seen in Table 5.1, the majority of participants were female (76.5%) approximately half were living in the USA, and a third in the UK. The mean age was 29 (ranging from 18 to 70).

A minimum sample size of 200 is often given as a rule-of-thumb to yield reliable results from SEM but this figure has been criticised as "conservative" and "simplistic" (Iacobucci, 2010). Anderson and Gerbing (1984) regard a sample size of 150 as sufficient, particularly when effects are large, the variables included in the model are reliable, and the models being tested are relatively simple. Although the latter two points do apply in the current study, the size of the effects are unknown, and so a sample size in excess of 200 was viewed as ideal, with 150-200 being seen as acceptable.

Table 5.1
Demographic Information

Demographic	Internet community sample (n = 234)
Mean age (SD)	28.84 (11.33)
Age range	18 - 70
Gender (% female)	76.50%
Country of residence:	
USA	52.56%
UK	30.34%
Europe (other)	4.70%
Asia	3.85%
Other	8.55%
Ethnic origin:	
White	88.89%
Asian	4.27%
Black	2.56%
Mixed	1.28%
Other ethnic group	2.99%
Treatment for a mental health problem (% yes)	50.86%

5.2.3 *Materials*

5.2.3.1 *Cognitive Fusion Questionnaire,* (CFQ; Gillanders et al., submitted) For details see Chapter IV, Section 4.2.2.

5.2.3.2 *Childhood Trauma Questionnaire* (CTQ; Bernstein, Fink, Handelsman, & Foote, 1994). The CTQ is a 28-item, self-report measure that retrospectively assesses a range of traumatic childhood experiences in five categories: sexual abuse, physical

abuse, emotional abuse, physical neglect and emotional neglect. The measure yields a total score or five subscales. Respondents indicate on a five-point scale (from *never true* to *very often true),* the accuracy of each of the items. Items include "When I was growing up people in my family said hurtful or insulting things to me" (an emotional abuse item), and "When I was growing up I had to wear dirty clothes" (a physical neglect item). The CTQ has very good psychometric properties, including a Cronbach's *α* figure of .91 for the measure as a whole (Scher, Stein, Asmundson, McCreary, & Forde, 2001), and test-retest reliability ranging from .79 to .86 over an average of four months (Bernstein et al.).

5.2.3.3 *Affect Intensity Measure - Negative Affectivity subscale* (AIM-NA; Larsen & Diener, 1987; Bryant, Yarnold, & Grimm, 1996). The 40-items of the AIM have been used as a whole scale and as various different subscales. Following Gratz (2006), a 12-item NA variable was created by combining the six negative intensity items (for example "When I do feel anxiety it is normally very strong"), and the six negative reactivity items (for example "When I talk in front of a group for the first time my voice gets shaky and my heart races"). This variable was selected because it most resembles the negative emotionality general genetic risk factor for PD identified by Kendler et al. (2008). The AIM has good psychometric properties, with Cronbach's *α* ranging from .87 to .90. Internal reliability figures for the negative intensity and reactivity subscales range from .66 to .72, the lower figures in all likelihood being a result of the small number of items in each subscale. Gratz (2006) did not report a Cronbach's *α* figure for the NA variable, but it is likely to be higher than the figures for the intensity and reactivity subscales.

5.2.3.4 Severity Indices of Personality Problems (SIPP; Verheul et al., 2008). The SIPP is a self-report measure of personality functioning. It consists of 118 items in five domains; self-control, identity integration, relational capacities, responsibility, and social concordance. Respondents indicate the extent to which they agree or disagree with each statement in relation to the previous 3 months. Items include 'I know exactly who I am and what I am worth' (an identity integration item), and 'I can work with people on a joint project in spite of personal differences' (a social concordance item). The SIPP has very good psychometric properties, with Cronbach's α for the five domains ranging from .74 to .79, with a mean of .77. Test-

retest reliability for the five domains, (over a period of two to three weeks), ranges from $r = .87$ to $r = .93$.

For the present study, an overall personality functioning variable was created by averaging scores across the five SIPP domains. The authors of the SIPP have not published data using the measure in this way, reporting that they have not explored the possibility of a single factor or higher order factor solution for the SIPP, with a community sample (Andrea, personal communication, 2011). Andrea notes, however, that the five domains appear to "hang together" much better in community samples than they do with PD samples, and speculates that the SIPP might well perform differently (and yield a different factor solution) with a non-PD sample.

Ideally, the factor structure of the SIPP with a community sample would be examined using EFA and CFA. Given the large number of items in the measure, this was not possible with the current sample, but the internal consistency of the overall SIPP variable and the inter-correlations between the five domains were assessed. This provided some preliminary evidence relevant to scoring the SIPP as a single variable.

5.2.3.5. *Brief Demographics Questionnaire*. Participants were asked to indicate their age, gender, country of residence, ethnic origin, and whether they had ever sought treatment for a mental health problem (see Appendix G).

5.2.4 *Procedure*

University of Southampton Psychology ethical approval and research governance approval were obtained. This online study provided potential participants with information about the nature of the study (on a webpage), prior to their giving consent to participate by clicking a button at the end of the information page (see Appendix I). Specifically, detailed warning was given that some questions were of a sensitive nature. Participants were advised to complete the study privately. It was emphasised that participants could end the study at any point, and then proceed to the final information page, which included not only the contact details of the researcher (an experienced clinical psychologist), but also contact details for UK and international support agencies.

5.2.5 *Analysis Plan*

Stage One of the analysis was designed to yield skew, kurtosis and internal consistency data for each of the study variables, as well as normative data. Stage Two involved Baron and Kenny's (1986) four-step mediational analysis (see Section 3.2.2.2) to assess CF as a mediating variable in the relationship between CT and adult personality functioning, and in a separate analysis, between NA and personality functioning. The bootstrapping procedure recommended by Preacher and Hayes (2004) was used to assess the significance of the indirect effect in the two mediation models (see Section 3.2.2.2 for more details).

Stage Three utilised SEM to test more complex mediation models involving both predictor variables. The fit of the various models to the data was evaluated using the following fit indices: χ^2/df ratio, comparative fit index (CFI), root mean square error of approximation (RMSEA), and Akaike's Information Criterion (AIC). An explanation for the use of these indices is given in Chapter IV (Section 4.2.4).

PASW/SPSS Statistics 19 software was used for all data analysis except SEM, which was conducted using AMOS 18 software.

5.3 Results

5.3.1 *Stage One: Data Screening and Preliminary Analyses*

Prior to conducting any statistical analyses, the data were examined in a number of ways. Of the 234 participants, 26 (11.11% of the sample) were found to have more than 10% of data missing from a questionnaire, but across participants no particular questionnaire was avoided. The data for these participants were excluded from subsequent analyses. The CFQ was the only questionnaire for which there were no examples of more than 10% of the data missing. All other missing data (less than 10% per questionnaire) was replaced by the sample mean (Tabachnick & Fidell, 2001).

A further four sets of data were excluded from the analysis as each of these participants had given the same response to all items on a measure, indicating that they were not completing the measures accurately. Mahalanobis distance values

identified two possible outliers, but examination of the data indicated that they appeared to be outliers because their scores on the CTQ were high (indicating substantial CT), whereas the distribution of sample responses in general was somewhat skewed towards low levels of trauma. Given that the data for these two participants were not contributing to this skew, and because it was considered important to include individuals who reported CT, their data were included in the analyses. This resulted in a final sample size of 204.

Table 5.2

Skew and kurtosis data for all study variables

Variable	skew	c.r.	kurtosis	c.r.
CFQ	.32	1.89	-.42	-1.23
SIPP	-.26	-1.50	-.88	-2.56
Neg AIM	.08	.49	-.72	-2.08
CTQ	1.48	8.66	1.99	5.86

Note. CFQ = Cognitive Fusion Questionnaire; SIPP = Severity Index of Personality Problems; Neg AIM = negative affectivity subscale of the Affect Intensity Measure; CTQ = Childhood Trauma Questionnaire; c.r. = critical ratio.

Table 5.2 shows skew and kurtosis data. There is no general agreement as to what constitutes problematic levels of skew and kurtosis (Byrne, 2010). Using the guidance given by West et al. (1995) that a figure of 7 or less is acceptable, it can be seen that none of the study variable distributions differ from normality sufficiently to be of concern.

Table 5.3 contains means and standard deviations for all variables. Interestingly, mean score on the CFQ (48.47) falls between mean CFQ scores for community samples (41.53) and a clinical sample (60.76) (Gillanders et al., submitted). This fits with the fact that approximately half the current sample identified themselves as having sought help for a mental health problem at some point. Table 5.3 also shows Cronbach's α scores. The CFQ, SIPP and CTQ had excellent internal consistency, and the negative affectivity subscale of the AIM had good internal consistency.

Table 5.3

Means, Standard Deviation, and Internal Reliability Values for All Study Variables

	Community Internet sample (n = 204)	
Variable	Mean (SD)	Cronbach's α
CFQ	48.47 (15.13)	.90
SIPP	5.01 (.77)	.97
Neg AIM	45.44 (9.79)	.82
CTQ	43.00 (17.43)	.94

Note. CFQ = Cognitive Fusion Questionnaire; SIPP = Severity Index of Personality Problems; Neg AIM = negative affectivity subscale of the Affect Intensity Measure; CTQ = Childhood Trauma Questionnaire.

Table 5.4 shows the inter-domain correlations for the five SIPP domains. These range from .42 to .81, with a median of .60. These figures are considerably higher than those reported by Verheul et al. (2008), with a PD sample, and provide support that the subscales of the SIPP are sufficiently highly related that an overall personality functioning variable can be created from them.

5.3.1.1 *Bivariate correlations*

Bivariate correlations between all study variables are reported in Table 5.5. All the variables correlated significantly with each other after the required p value was adjusted for multiple comparisons (p = .05/4 = .0125). It should be noted that because the SIPP is scored so that a higher score represents more adaptive personality functioning, whereas all the other measures are scored so that higher score indicates poorer functioning or a greater level of problem, the SIPP has inverse relationships with all other study variables.

Table 5.4

Inter-Correlations Between SIPP Subscales

	SIPP domains				
	Self cont	Social Conc	Identity Int	Relation	Responsib
Self cont	-	.70*	.81*	.58*	.67*
Social Conc		-	.61*	.57*	.52*
Identity Int			-	.76*	.59*
Relation				-	.42*
Responsib					-

Note. Self cont = Self control; Social Conc = Social Concordance; Indentity int = Identity integration; Relation = Relational Capacities; Responsib = Responsibility. All are subscales of the Severity Index of Personality Problems.
*≤ .005 (adjusted for multiple comparisons).

Table 5.5

Bivariate Correlations Between Study Variables

	CFQ	SIPP	NegAIM	CTQ
CFQ	-	-.77***	.51***	.35***
SIPP		-	-.39***	-.33***
NegAIM			-	.18**
CTQ				-

Note. CFQ = Cognitive Fusion Questionnaire; SIPP = Severity Index of Personality Problems; Neg AIM = negative affectivity subscale of the Affect Intensity Measure; CTQ = Childhood Trauma Questionnaire.
** = significant at p. < .01; *** = significant at p. < .001.

5.3.2 *Stage Two: Mediational Analyses*

CF was tested as a mediating variable in the relationships between NA and personality functioning, and separately, between CT and personality functioning. These analyses followed Baron and Kenny's (1986) causal step approach (see Section

3.2.2.2). The mediation models are displayed graphically in Figure 5.1. Focussing firstly on NA as a predictor, the four mediation steps were satisfied as follows:

1. NA significantly predicted personality functioning: $t(202) = -6.06$, $p < .001$, $\beta = -.39$, $R^2 = .15$.
2. NA significantly predicted CF: $t(202) = 8.35$, $p < .0001$, $\beta = .51$, $R^2 = .26$.
3. CF significantly predicted personality functioning: $t(202) = -17.19$, $p < .001$, $\beta = -.77$, $R^2 = .59$.
4. When both NA and CF (the mediator) were included in the model, NA no longer significantly predicted personality functioning ($t(201) = -.04$, $p = .97$, $\beta = -.002$), whilst CF still significantly predicted it ($t(201) = 14.77$, $p < .001$, $\beta = -.77$). R^2 for the model was .59, with NA accounting for no variance in the DV other than indirectly via CF.

Focussing on CT as a predictor, the four steps were satisfied as follows:

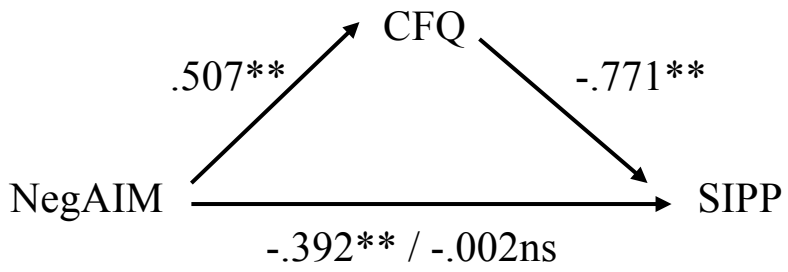
1. CT significantly predicted personality functioning: $t(202) = -4.94$, $p < .001$, $\beta = -.33$, $R^2 = .11$.
2. CT significantly predicted CF: $t(202) = 5.25$, $p < .0001$, $\beta = -3.28$, $R^2 = .12$.
3. CF significantly predicted personality functioning: $t(202) = -17.19$, $p < .001$, $\beta = -.77$, $R^2 = .59$.
4. When both CT and CF (the mediator) were included in the model, CT no longer significantly predicted personality functioning ($t(201) = -1.46$, $p = .15$, $\beta = -.07$), whilst CF still significantly predicted it ($t(201) = 15.67$, $p < .001$, $\beta = -.75$). R^2 for the model was .60, with CT accounting for no variance in the DV other than indirectly via CF.

In each case, the significance of the indirect effect (via CF) was tested using the Preacher and Hayes (2004) bootstrapping method (see Table 5.6). Confidence intervals that do not include zero indicate a significant indirect effect, which was the case with both mediated relationships. Overall, these findings suggest that CF may play an important role in the relationships between risk factors and problematic personality functioning. However, models of PD development such as Linehan's (1993) biosocial model of BPD suggest that these kinds of risk factors are likely to be related, so the final stage of analysis involved using SEM to test more complex mediational models involving both NA and CT.

Figure 5.1

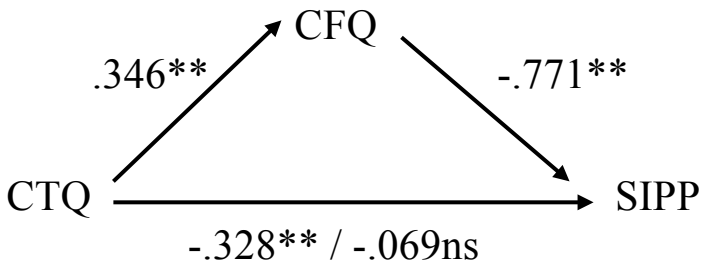
Mediation Models Involving Each Risk Factor:

1. NA



Note. Neg AIM = negative affectivity subscale of the Affect Intensity Measure; CFQ = Cognitive Fusion Questionnaire; SIPP = Severity Index of Personality Problems
ns = non-significant standardised regression coefficient; ** standardised regression coefficient significant $p < .001$

2. CT



Note. CTQ = Childhood Trauma Questionnaire; CFQ = Cognitive Fusion Questionnaire; SIPP = Severity Index of Personality Problems
ns = non-significant standardised regression coefficient; ** standardised regression coefficient significant $p < .001$

Table 5.6

Tests of CF as a Mediator of Significant Associations between Independent and Dependant Variables

	Indirect Effects		BCa 95% CI	
Variables	Point Estimate	SE	Lower	Upper
CFQ mediating the impact of Neg AIM	-.0305	.0042	-.0395	-.0225
CFQ mediating the impact of CTQ	-.0114	.0023	-.0160	-.0070

Note. BCa Cl = bias corrected and accelerated confidence intervals, derived from 2,000 bootstrap samples of the data. Neg AIM = negative affectivity subscale of the Affect Intensity Measure; CTQ = Childhood Trauma Questionnaire; CFQ = Cognitive Fusion Questionnaire.

5.3.3 *Stage Three. SEM Mediational Models*

The final analysis stage involved testing the fit of a hypothesised full mediation model (Figure 5.2, Model B), to the data, and comparing the fit of this model with other, theoretically justifiable models. The four models tested (see Figure 5.2) were:

Model A – fully mediated model with the two risk factors (NA and CT) unrelated; fusion mediating the relationships between risk factors and personality functioning

Model B – fully mediated model, as in Model A, but with the risk factors covarying

Model C – fully mediated model with the risk factors covarying, with personality functioning acting as the mediator in the relationships between risk factors and fusion

Model D – partially mediated model with the two risk factors covarying, with a direct path between CT and personality functioning, as well as the indirect path via fusion. The direct path from NA to personality functioning was not included because the indirect effect of fusion on this path was so substantial that adding the direct path would be highly unlikely to improve the model fit.

It can be seen from Figure 5.2. that all the regression coefficients in all four models were significant and signed in expected directions, except for the direct path

between CT and personality functioning, which was non-significant (Model D). This lack of significance suggests that the path is redundant and can be removed. Nonetheless, the fit of all four models was examined. In a well-fitting model, χ^2 /df should be less than 2, the CFI value should be above .95, and the RMSEA value should be less than .05. The AIC is a fit index that takes parsimony into consideration, with a lower value indicating a better fit.

Table 5.7.
Fit Indices for Four Mediation Models

Model	Fit index					
	χ^2	*df*	χ^2 /*df*	CFI	RMSEA	AIC
A	8.83	3	2.94	.98	.10	53.06
B	2.13	2	1.07	1.0	.02	52.68
C	30.63	2	15.31	.89	.27	81.17
D	0	1	0	1.0	0	56.86

Note. χ^2 = minimum fit function chi-squared; df = degrees of freedom; CFI = comparative fit index; RMSEA = root mean square error of approximation; AIC = Akaike's Information Criterion

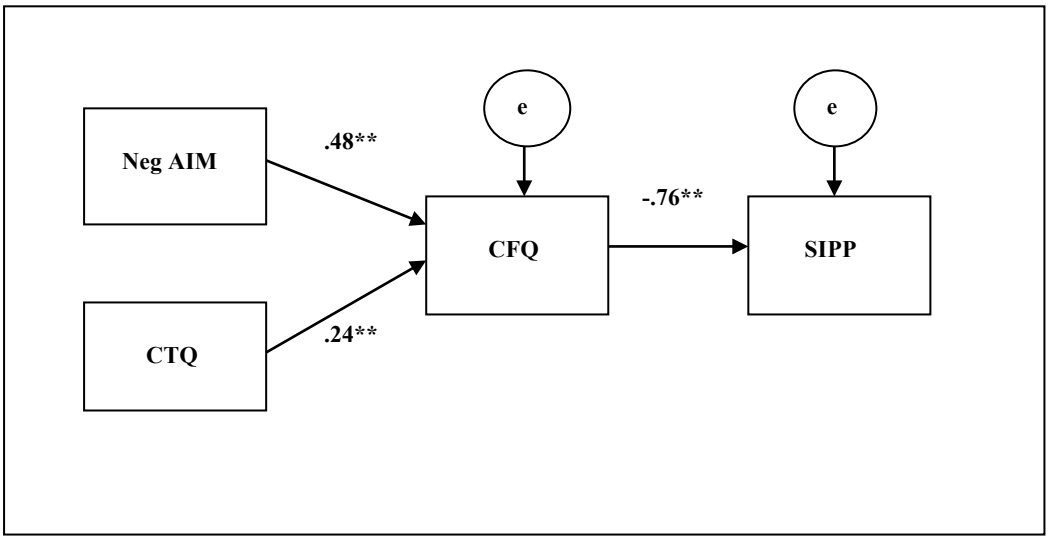
Table 5.7 shows that the that Models A and C did not fit the data well in terms of the various fit indices: these were therefore rejected in favour of the other two models. Models B and D, the full and partial mediation models both fitted the data very well. However, it has already been noted that the direct path from the CTQ to the SIPP was non-significant. A χ^2-difference test was performed between models B and D as follows:

Model B χ^2 = 2.13(2) - Model D: χ^2 = .001(1) = χ^2 -difference = 2.129(1).

The critical value for χ^2 with 1 degree of freedom is 5.99 ($p < .05$), and as 2.129 is less than this critical value, it was concluded that adding the direct path from CT to personality functioning did not improve the fit of the model to the data. The AIC values also indicated that Model B, the fully mediated model, represented the best fitting, most parsimonious model. In this model, CF predicted 59% of the variance in personality functioning.

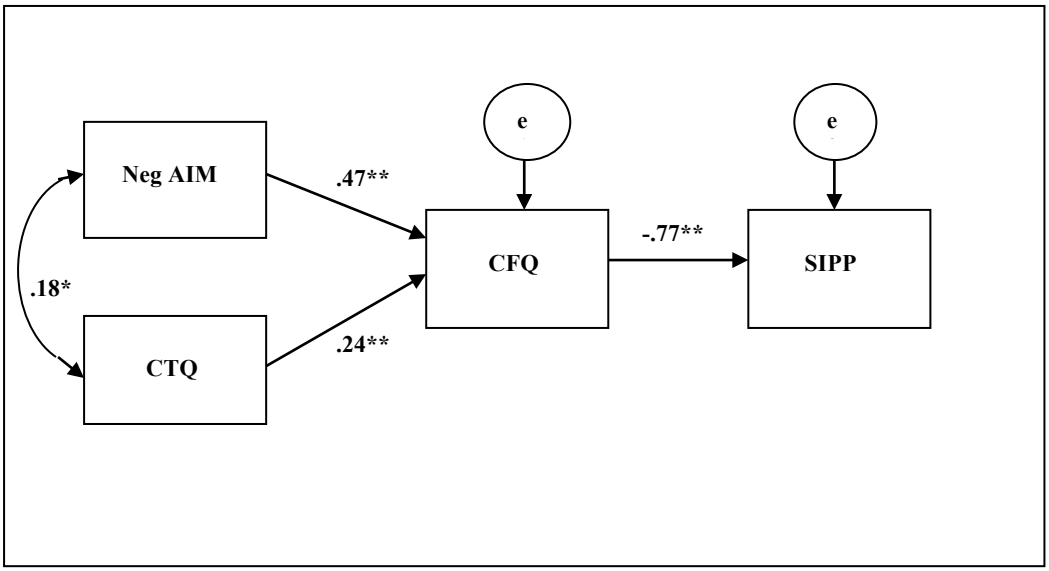
Figure 5.2

SEM Path Diagrams: Model A



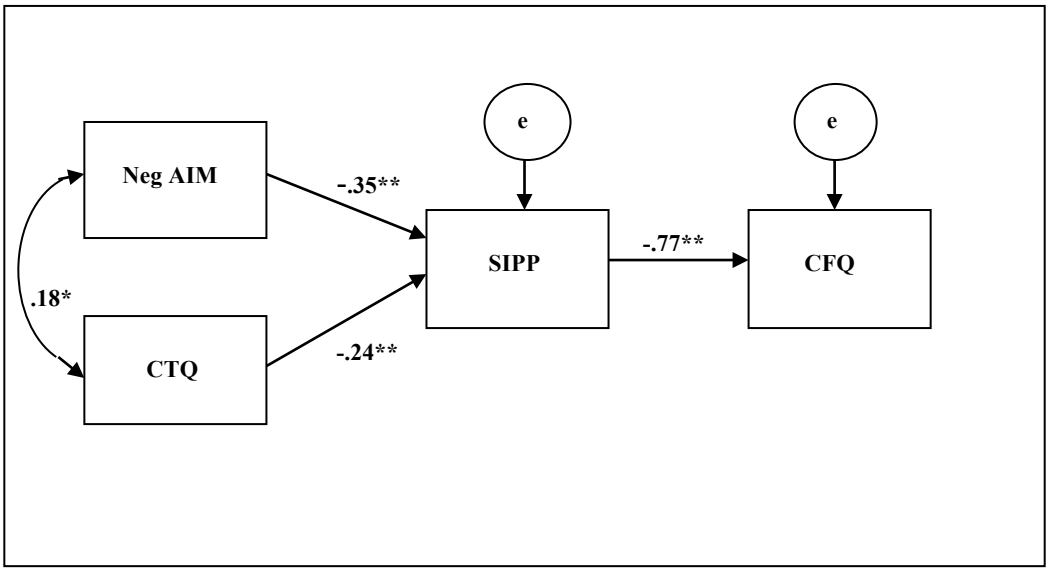
Note. CFQ = Cognitive Fusion Questionnaire; SIPP = Severity Index of Personality Problems; Neg AIM = negative affectivity subscale of the Affect Intensity Measure; CTQ = Childhood Trauma Questionnaire; e1 and e2 denote error terms.
** represents standardised regression weights significant at p. < .001.

SEM Path Diagrams: Model B



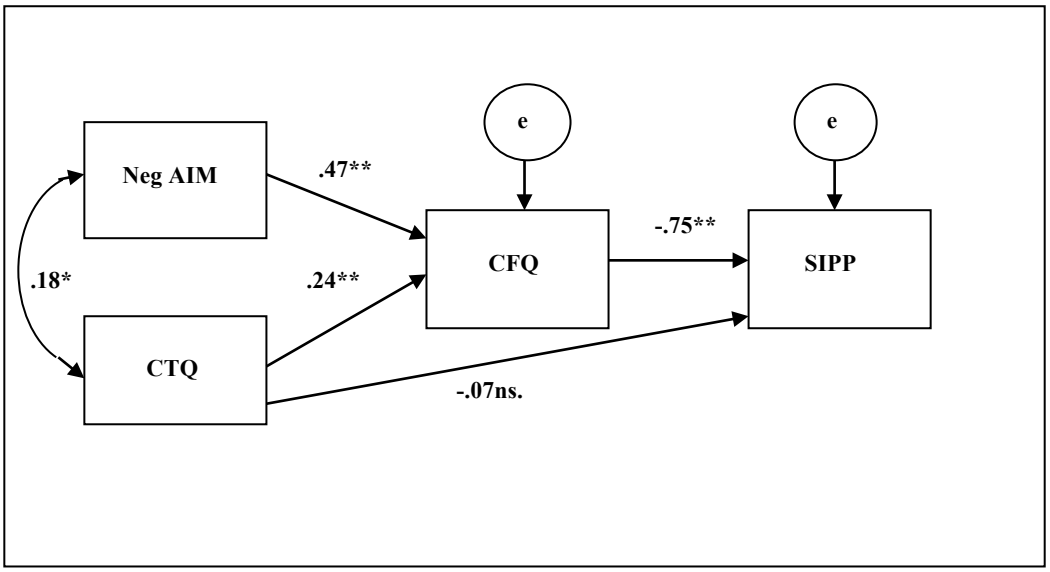
Note. CFQ = Cognitive Fusion Questionnaire; SIPP = Severity Index of Personality Problems; Neg AIM = negative affectivity subscale of the Affect Intensity Measure; CTQ = Childhood Trauma Questionnaire; e1 and e2 denote error terms.
** represents standardised regression weights significant at p. < .001. * represents covariance significant at p. < .05.

SEM Path Diagrams: Model C



Note. CFQ = Cognitive Fusion Questionnaire; SIPP = Severity Index of Personality Problems; Neg AIM = negative affectivity subscale of the Affect Intensity Measure; CTQ = Childhood Trauma Questionnaire; e1 and e2 denote error terms.
** represents standardised regression weights significant at p. < .001. * represents covariance significant at p. < .05.

SEM Path Diagrams: Model D



Note. CFQ = Cognitive Fusion Questionnaire; SIPP = Severity Index of Personality Problems; Neg AIM = negative affectivity subscale of the Affect Intensity Measure; CTQ = Childhood Trauma Questionnaire; e1 and e2 denote error terms.
** represents standardised regression weights significant at p. < .001. ns. represents non-significant standardised regression weight. * represents covariance significant at p. < .05.

5.4 Discussion

5.4.1 *Study Findings*

Previous research has identified both NA and CT as risk factors for poor personality functioning in adult life. It is, however, important to establish the processes through which these risk factors impact personality functioning, to develop theoretical understanding and to guide the development of more effective psychosocial interventions (the modelling phase of treatment development, outlined by Campbell et al., 2000).

The ACT model of psychopathology suggests that CF is implicated in the development of psychological difficulties across diagnoses. However, the role of fusion with personality functioning has not been examined empirically to date. This study therefore aimed to test CF as a mediating variable in the relationships between the two risk factors and personality functioning. It was predicated that CF would fully mediate both relationships.

5.4.1.1. *Mediational Analyses*

As predicted, CF was shown to fully mediate the relationships between the two risk factors and personality functioning. Causality cannot be inferred from these results owing to the cross-sectional design of the study. However, the results are consistent with the account that (in relation to NA), it is individuals' fused relationship with their private experiences—rather than propensity to experience negative emotions intensely and reactively per se—that leads to them developing personality problems in adult life. Similarly, these results are consistent with the view that for people who have experienced CT, their relationship to their thoughts and memories influence whether they go on to develop problematic personality functioning.

It has been hypothesised that genetic and environmental risk factors for personality problems do not influence personality functioning in isolation, but rather that such risk factors transact with each other, at least in relation to BPD (Linehan, 1993). When both NA and CT were included in an SEM mediation model together, it was found that that they were indeed significantly related. The model that stipulated

that they were independent of each other fitted the data poorly in comparison to the model where they co-varied. This finding is supportive of a general interactive risk model in relation to poor personality functioning across diagnoses.

When both risk factors were included in SEM models, it was found that a model where CF fully mediated the relationships between the risk factors and personality functioning was the best fitting and most parsimonious, consistent with the results from the single-risk factor mediation analyses. This model fitted the data very well, accounting for 59% of the variance in the DV.

5.4.1.2. *Performance of the CFQ*

The results of this study provide further support for the reliability, validity, and utility of the CFQ. With this community-based sample, approximately 50% of whom had sought help for a psychological problem, the CFQ demonstrated excellent internal reliability. It also related in a theoretically consistent manner to the other study variables. It was also the only study questionnaire without a single instance of a participant leaving more than 10% of the items incomplete, suggesting that participants found it straightforward to complete. Overall, the CFQ showed itself to be an excellent measure for the purpose of this study.

5.4.1.3. *Performance of the SIPP Composite Variable*

In order to test relatively simple models, a composite variable was created from the SIPP subscales. The measure had not been used in this way before. Verheul et al. (2008) selected a five-factor SIPP model on the basis of CFA, but did not report having tested a single factor or higher order factor model. Unfortunately, the current sample was not large enough to conduct factor analysis, but it was possible to conduct three forms of analyses relevant to this issue. Cronbach's alpha was calculated for the composite version of the measure, and it was found to be very high (.97), indicating excellent internal consistency. Cronbach's alpha is influenced by the number of items in a measure, and as the SIPP has 118 items, this figure could have been thus inflated, but other published personality problem measures with similar numbers of items have much lower Cronbach's alpha values. For example, the mean alpha across several samples for the PDQ (Hyler et al., 1988), which contains 99 items, is .71. Clearly a large number of items is not sufficient to produce a high alpha value. The apparently excellent internal consistency of the SIPP variable indicates

that the items are highly intercorrelated, and could be consistent with them measuring the same construct.

The intercorrelations between the five SIPP subscales were also calculated. They ranged from .42 to .81, with a median of .60. These figures are substantially higher than those reported by Verheul et al. from their clinical sample (.27 to .60, with a median of .40). This high subscale intercorrelation suggests that, for this sample at least, they are sufficiently related for it to be acceptable to combine them to produce an overall indicator of personality functioning. Certainly, there are published measures with comparable subscale intercorrelations where the measure is scored both in terms of its individual subscales and as a single measure. One example is the Psychological Inflexibility in Pain Scale (PIPS; Wicksell, Renofalt, Olsson, Bond, & Melin, 2008), the authors of which reported a two-factor structure with a correlation of .46 between the two resultant subscales. The measure is scored either as a single measure or as two subscales.

Finally on this matter, the SIPP composite variable related in theoretically consistent ways with the other study variables. That said, the high correlation between this variable and the CFQ raises questions about the relationship between ACT-relevant processes such as fusion, and personality. Hayes et al. (2006) argue that psychological inflexibility (to which fusion contributes) is highly predictive of—and represents a general vulnerability for—psychopathology across diagnostic categories. From this perspective, it is not surprising that score on the CFQ should be seen to so strongly predict score on the SIPP in the mediation model yielded by this study. However, another possible explanation for this finding is that psychological inflexibility is actually personality by another name. Further research (such as testing the predictive utility of the CFQ and the AAQII above and beyond measures of personality) is required to clarify this issue. However, the fact that there was very poor model fit when the SIPP was tested as a mediator in relation to the CFQ as a DV (Model C), might suggest that the processes being measured by these two questionnaires are not equivalent.

Overall, these findings support scoring the SIPP as a global measure of personality functioning, and suggest that the resultant measure is both reliable and valid.

5.4.2 *Methodological Limitations*

A number of limitations as well as strengths should be taken into account when considering the results of this study. In terms of strengths, the sample was large enough to ensure confidence in the results of the analyses. It included participants with a wide range of ages, unlike many published studies based on young, student samples, and the sample was international. On the other hand, there was bias in favour of female participants, and the sample was predominantly Caucasian. Although the use of internet-based research websites ensured a large and geographically varied sample, this methodological choice also resulted in a self-selecting, opportunity sample, and provided no means of verifying demographic and other data independent of the information given by participants themselves. Caution should therefore be exercised when generalising from these results.

The cross-sectional design of the study limits the conclusions that can be drawn from the findings. Despite both IVs having been identified as risk factors for poor personality functioning in other research including longitudinal studies, causality cannot be inferred from the current study, and these mediational models should be tested in studies that measure the relevant variables across time. Obviously, for ethical reasons, these particular IVs cannot be manipulated, although 'natural experiments' such as the work of Rutter and colleagues examining the development of adopted children from the UK and Romania (Rutter, 2004) can go some way to address these kinds of methodological limitations. There are uncontrolled, prospective studies that follow children into adulthood and track the impact of CT (for example) on personality development (e.g. Johnson et al., 1999) but they are extremely expensive and by their nature, take many years to complete. It is unrealistic to consider these as viable methodological options specifically for testing potential mediating variables.

However, as a body of evidence develops regarding risk factors for poor personality functioning (to which this study has contributed), it is possible that interest will grow in relevant mediating processes, and as a result, measures of potential mediating variables may in future be included in large-scale, longitudinal studies. Another, more realistic way to examine the mediating role of CF (and other possible mediators and moderators) in relationship to personality functioning would

be to include them as process variables in RCTs, to test whether they mediate personality functioning outcomes. This will be discussed further in Section 5.4.3.

Although the findings of this study represent an important addition to the scientific understanding of personality functioning, other temperamental predictors in addition to NA—such as impulsive aggression and inhibition—have been identified (Kendler et al., 2008). It is a limitation of this study that only NA was included in the models tested, and further research should build on this study by testing more complex models.

Finally, in terms of methodological issues, the dependence on self-report measures in this study, although in keeping with much of the relevant published research, is a limitation. It is possible to gather data about CT in particular, through other means such as examining childhood health and social care records, and through the use of interview. However, each of these methods is associated with other potential problems, such as the lack of participant anonymity perhaps resulting in poorer recruitment, and possible biases in the personal disclosure and official reporting of childhood trauma.

5.4.3 *Implications and Future Research*

The results of this study are important for a number of reasons. They provide the first empirical evidence that CF is related to poor personality functioning. The findings are consistent with clinical observation, and also provide support for the ACT model of psychopathology, which, although conceptualised as transdiagnostic, has never been investigated in relation to personality functioning across diagnoses. Given that risk factors have been identified that appear to negatively impact personality functioning across many PD diagnostic categories, it makes sense to investigate intermediary processes that are also thought to act transdiagnostically. Of course, the picture is more complicated than this, as studies investigating risk factors for personality difficulties have also identified factors that appear to predict specific personality problem presentations. CF, despite accounting for a substantial proportion of the variance in personality functioning in the models tested in this study, will not be the only mediating variable of relevance in this complex situation. For example, it may be that whilst fusion mediates the action of general predictors, other processes may mediate the impact of more specific risk factors. Future research should include

a systematic examination of an array of theoretically meaningful potential mediators and moderators, an issue that is returned to below in relation to DBT and EA.

Fusion has not previously been tested as a mediating variable in relation to personality problems, and in fact due to measurement issues (see Chapter IV for more details), it has rarely been examined as a mediating variable in relation to any psychological difficulties. This study has therefore made a significant scientific contribution both in terms of examining an aspect of the ACT model empirically, as well as increasing knowledge about influences on the development of personality difficulties. However, these findings are based on a sample that includes people across the spectrum of personality functioning, and it would be useful to test these mediation models with a clinical sample of people with personality difficulties.

These results have important clinical as well as theoretical implications. Given that adults with personality problems can neither change their histories nor their genetic make-up, it is vital to identify intermediary processes that impact the development and maintenance of such problems, which can potentially be addressed therapeutically. CF plays an important part in the ACT model, and defusion exercises are included in all ACT treatment protocols, so it might be assumed that ACT would be helpful in treating poor personality functioning. However, ACT remains virtually untested in relation to personality difficulties (see Chapter II for a review of the few relevant trials), and PDs tend to be substantially more difficult to treat than Axis I disorders (Bender et al., 2001). Whilst the findings of the current study indicate the relevance of CF to personality functioning, it does not necessarily follow that ACT will prove to be a safe and effective treatment for people with poor personality functioning. Careful treatment development work is needed to explore this possibility.

DBT is already well established as a psychosocial intervention for BPD. It is possible that defusion is one of the mechanisms of change in DBT with this patient group, and this possibility should be tested empirically. In their 2006 theoretical paper outlining possible mechanisms of change in DBT for BPD, Lynch, Chapman, Rosenthal, Kuo, and Linehan consider mindfulness and the reduction of literal belief in rules amongst the likely mechanisms of change, both of which could be understood in terms of fusion/defusion. Additionally, DBT treatment strategies such as irreverence could also be viewed as supporting defusion. Lynch et al. in fact suggest several possible mechanisms of change, thus pointing to the need to test multiple

potential mediators and the relationships between these mediators, both in relation to risk factors and treatment outcomes.

It has been argued (Greco et al., 2008) that CF leads to EA, although there is in fact a lack of clarity about how the various ACT-relevant variables relate to each other (see Chapter II, Section 2.2.1.2). It would therefore be theoretically important to test both CF and EA as mediators. As discussed in Chapter IV, the AAQ (Hayes et al., 2004), which has commonly been used to measure EA, includes items that address psychological flexibility rather than just those relating to EA, and includes fusion-related items. Consequently, the AAQ may not be the best tool for examining the relationship between CF and EA and their respective roles in relation to personality functioning. A more precise and/or non-questionnaire based measure of EA should be developed before these processes are tested together as mediating variables, in order to yield data that will help refine the ACT model. Given that avoidant behaviours are so commonly observed amongst people with poor personality functioning, a measure of the behavioural aspects of EA might be particularly useful.

Finally, and somewhat at a tangent to the focus of the current study, the study findings provide initial support for the use of a composite SIPP variable that gives a measure of overall personality functioning. Future research in the form of EFA and CFA with sufficiently large community and personality difficulty samples is needed to continue the examination of the factor structure and psychometric performance of this variable.

5.4.4 *Summary*

The results of this study provide the first empirical evidence that CF may play a significant mediating role in the relationships between risk factors for personality problems and actual personality functioning in adulthood. This has important theoretical implications, in terms of increasing understanding of personality difficulties, as well as providing support for the ACT model of psychopathology. These findings are also important clinically, in that they suggest that if fusion can be effectively addressed through psychotherapeutic interventions, this might have a positive impact on outcome in the treatment of personality problems.

A number of directions for future research suggest themselves from the findings of the present study. In terms of theory, it is important to understand more about fusion and how it relates to other processes and behaviours, particularly those

relevant to ACT. Given the limitations of the self-report measures of ACT-relevant processes currently, a fruitful first step might be to test fusion in relation to a behavioural measure relevant to personality problems. Clinically, and in relation to poor personality functioning, the next logical step, based on the models of treatment development outlined in Chapter III, is a small scale, uncontrolled treatment development trial of ACT for poor personality functioning.

CHAPTER VI

Study 3: An Investigation of the Behavioural Correlates of CF

6.1 Introduction

Research focusing on the behavioural consequences of CF has been relatively limited to date, to a large extent owing to the lack of an appropriate measure of the construct. Study 1 and Study 2 of this thesis reported on the validation of a new self-report measure of CF and used it to demonstrate how CF mediated aspects of poor personality functioning. An additional benefit of the development of the CFQ is that it is now possible to investigate empirically the consequences of CF in a variety of novel ways. The current study assessed the behavioural consequences of fusion, specifically in relation to EA, because these two variables together play a crucial role in the ACT model of psychopathology (Hayes et al., 1999). Behaviours that serve to avoid distressing thoughts and emotions, such as DSH and binge eating, are commonly observed with people with PD diagnoses, and appear to have a detrimental effect on psychosocial functioning. Indeed, Kingston et al. (2010) found that report of such behaviours was strongly related to known risk factors for PD. Therefore, developing methodologies that enhance the investigation of CF and EA is likely to yield clinically relevant as well as theoretically important findings.

6.1.1 *Laboratory-Based Investigation of CF*

Recent guidance on the evaluation of psychosocial interventions has emphasised that, in addition to establishing the efficacy of interventions, it is necessary to test the theories and models on which such interventions are based (David & Montgomery, 2011; Lohr, 2011). Levin et al., 2012, have argued that laboratory-based research can make an important contribution to this process, because it provides an opportunity to isolate and manipulate variables in a more precise and controlled manner than would be possible when testing complex psychotherapy treatment protocols in clinical settings.

As discussed in Chapter IV, the lack of a psychometrically sound, widely applicable measure of CF has hindered empirical examination of the construct, including in the laboratory. The few relevant, laboratory-based publications are component studies, each experimentally testing a standalone, brief defusion intervention, usually in relation to aversive stimuli such as negative self-statements (see Chapter II, Section 2.2.1.2 for a review). There are no published studies specifically designed to investigate fusion (rather than defusion interventions).

Theoretically, it is possible to fuse with positive or neutral stimuli (Blackledge, 2007), but the few published experimental investigations of CF have all focused primarily on defusion in relation to negative or aversive material, not only because there is an assumption that fusion with negative material is more problematic than fusion with positive or neutral material (e.g. Healy, et al., 2008), but also because there appears to be an untested—and usually unarticulated—assumption that we fuse more readily with negative private experiences than with positive. To date, no studies designed to test this assumption have been published. More surprisingly still, given the central role of these processes in the ACT model, few studies—laboratory-based or otherwise—have investigated the relationship between CF and EA.

6.1.2 *CF and EA*

Although examining the impact of components of complex psychotherapies is a central aspect of model testing (Campbell et al., 2000; Levin et al., 2012), evaluating whether that model's variables relate to each other in predicted ways is also important. Few published studies have done this in relation to CF and variables relevant to EA. As outlined in Chapter II, Section 2.2.1.2, EA is a process central to the ACT model that has been defined as "the phenomenon that occurs when a person is unwilling to remain in contact with particular private experiences (e.g., bodily sensations, emotions, thoughts, memories, behavioural predispositions) and takes steps to alter the form or frequency of these events and the contexts that occasion them" (Hayes et al., 1996). ACT theory suggests that CF should be strongly related to EA, with some authors arguing that CF leads to EA, in that there would be no need to attempt to avoid a private experience such as a judgemental thought about oneself,

unless one was fused with it (Greco, et al., 2008; Pistorello et al., 2000). Efforts to avoid private experiences and the situations that might trigger them are commonly observed in psychotherapeutic work with people with PD diagnoses, and clinical observation indeed suggests that fusion with uncomfortable private experiences leads to EA.

In their test of a brief defusion exercise with a student sample, Healy, et al, (2008) found that defusion was associated with greater participant willingness to read and think about negative self-statements. This finding could be seen as supporting the ACT model, given that willingness is viewed in ACT as the opposite of EA, in the same way that fusion and defusion are viewed as being inversely related. However, willingness was indexed using a self-report rating scale from 0 to 100, rather than a psychometrically tested instrument or a behavioural measure of willingness, so caution should be used when drawing conclusions from this finding. Watson et al. (2010), in their investigation of defusion in relation to obsessional compulsive symptomology, found that a brief defusion practice had no impact on EA, as measured by the AAQ (Hayes et al., 2004). As will be discussed in Section 6.1.3, the use of the AAQ as a measure of EA is problematic. Thus, although it is possible that the lack of impact of this defusion exercise on AAQ score indicates that fusion and EA are not related as outlined in the ACT model, it is equally possible that the measures used in the study do not adequately address the constructs in question. Additionally, it could also be that CF and EA are related, but that the defusion exercise used in this study did not impact CF sufficiently to affect those relationships.

Finally, Wicksell et al. (2008) labelled the two subscales of their measure of psychological inflexibility in physical pain, ‘cognitive fusion’ and ‘avoidance of pain’, and found that they were moderately, positively correlated. However, given that this measure is designed for use with a physical pain sample and its items refer only to pain, it cannot be assumed that the same relationship between the two variables would be found in the general population or mental health samples, in relation to private experiences other than pain.

In summary, the published literature examining the relationship between CF and EA is limited, both in terms of the number and quality of studies, although both ACT theory and some limited relevant data suggest that CF and EA are positively related. Research designed to examine the relationship between the two constructs without employing clinical-type interventions, as well as research utilising better

quality, more appropriate measures, would aid testing of this aspect of the ACT model. Given the role that these processes are hypothesised to play in psychopathology, increasing understanding of the relationship between CF and EA is important for both clinical and theoretical reasons. Self-report measurement of CF has significantly improved with the development of the CFQ, but the existing measure of EA remains problematic.

6.1.3 *Measuring EA*

The AAQ (Hayes, et al., 2004) has been used as a measure of EA extensively in the ACT literature. However, as Hayes and colleagues (2006, p. 10) point out, although the AAQ was originally described as a measure of EA, it in fact " measures the degree to which an individual fuses with thoughts, avoids feelings, and is unable to act in the presence of difficult private events". It therefore measures a much broader construct than EA, and the revised version of the measure, the AAQII, (Bond et al., 2011), is now described as a measure of psychological inflexibility. This revision of the measure and the construct it is designed to measure is particularly problematic, given that the AAQ/AAQII is the main ACT process measure. There are other measures, such as the Emotion Control Questionnaire (Roger & Najarian, 1989) and the State Emotion-Regulation Questionnaire (Kashdan & Steger, 2006) that include items that address EA, but again, both are measures of broader constructs.

Given that no precise self-report measure of EA for use with the general population or mental health samples is currently available, the use of an alternative form of measurement is indicated. One possibility is a measure of behavioural avoidance. This would be particularly helpful in terms of clinical relevance, because such avoidance is frequently observed with PD patients. In terms of the definition of EA quoted in Section 6.1.2, such a measure would be assessing the action taken to alter the form or frequency of private experiences and the contexts that occasion them.

In clinical settings, a simple and effective measure, the behavioural approach test (BAT), is commonly used to assess behavioural approach/avoidance responses to aversive stimuli in relation to anxiety disorders (Anthony & Swinson, 2000). This task measures how near to the stimulus the patient is willing to approach, or how long they are willing to stay in contact with the stimulus. A particular strength of the BAT

is that it directly measures approach/avoidance, and therefore the results are straightforward to interpret.

In non-clinical research settings, several computer-based measures of behavioural approach/avoidance responses have been developed (see Krieglmeyer & Deutsch, 2010 for a review). These measures are easy to administer in a controlled manner using a computer, but they tend to measure approach/avoidance behaviours in a less direct manner than the BAT. For example, the manikin task employs an image of a manikin on a computer monitor that the individual can move towards or away from the stimulus, rather than, for example, the participant actually moving away from the stimulus or being able to avoid it by clearing the screen. Other tasks make use of the affective Simon effect (Simon & Rudell, 1967; de Houwer, Crombez, Baeyens, & Hermans, 2001), where the participant is instructed to respond to a non-affective cue such as the grammatical category of word stimuli (for example noun or verb), but a supposedly irrelevant, affective cue, (positive or negative word meaning), impacts the participant's performance on the task. These tasks yield reaction time data, from which the extent of approach or avoidance behaviour is inferred.

Using yet another computer-based task design, Cochrane, Barnes-Holmes, Barnes-Holmes, Stewart, and Luciano (2007) found that participants who scored highly on the AAQ (Hayes et al., 2004)—which was intended as a measure of EA by the authors—took longer than low-scoring participants to react in a task that required them to choose whether to view aversive or neutral images. High AAQ scoring participants also—based on the author's interpretation of *event-related potential* (ERP) data—possibly might have been engaging in more verbal activity (which the authors interpreted as verbally-based avoidance strategies) than those with low AAQ scores. Although these findings are generally supportive of the ACT model, they are based on the interpretation of reaction time data, in a study in which the AAQ plays a central role, and with a very small sample size (six participants per group in the ERP experiment). This approach can therefore only be seen as raising interesting possibilities for further research rather than as resulting in a valid and reliable behavioural measure of avoidance.

For the purposes of this study, I developed a measure of the behavioural aspect of EA that drew on the strengths of both the BAT and the computer-based tasks discussed above. The measure and its administration are described in detail in Sections 6.2.2.1 and 6.2.3. In brief, a simple, computer-based task allowed

participants to directly avoid prolonged contact with stimuli—in this case, negatively and positively valenced self-referential adjectives—by removing them from the monitor. The length of time each participant left the stimuli on the monitor was used as the measure of behavioural avoidance. Although the majority of defusion-focused studies have used only negatively valenced word stimuli, it was decided to additionally include positively valenced stimuli in order more flexibly to examine assumptions in the ACT literature about CF.

6.1.4 *Study Design, Aims, and Hypotheses*

The purpose of the present study was to develop a measure of the behavioural aspect of EA and to use it in conjunction with the CFQ to examine the relationship between CF (the IV) and EA (the DV). Based on ACT theory, it was expected that CF would positively correlate with EA (view time for stimuli), at least in relation to negative stimuli. However, if participants fully understood the purpose of the study, this knowledge could impact their behaviour, rendering the measure invalid. It was therefore necessary to represent the experiment as a test of memory, in order to disguise its true purpose (see Section 6.2.3 for the details of this aspect of the task). Thus, participants, having been told that there would be a memory test later, were presented with the stimuli, one at a time, which they viewed for as long as they wished. A stimulus recognition memory test duly followed and recognition data were collected and analysed on an exploratory basis rather than to test specific hypotheses, because any relationship between CF and recognition might be confounded by stimulus view time.

Although the processes and behaviours of interest in this study are observed in psychotherapy with people with PD diagnoses, it was decided to recruit a non-clinical, student sample in the first instance. This decision was made for both ethical reasons (not wishing to trial a methodology that features negative, self-referential stimuli with vulnerable participants), and practical reasons (the difficulty of recruiting a clinical sample).

The hypotheses were as follows:

1. CF, as measured by the CFQ, will be negatively correlated with view time for negative words.

2. CF, as measured by the CFQ, will be positively correlated with view time for positive words.

6.2 Method

6.2.1 *Participants*

An opportunity sample of fluent English speaking volunteer undergraduate and postgraduate psychology students at the University of Southampton (n = 40) was recruited (see Table 6.1 for demographic information). Undergraduate participants earned course credits for their participation.

It was difficult to calculate the required sample size for the study, because there were no published experiments using the same measure of behavioural avoidance. A required sample size of 28 was estimated, assuming statistical power of .8, a large effect size (r = .5), and using α = .05. Allowing for the likelihood of some invalid or incomplete datasets, and/or for the possibility of a smaller effect size, it was decided to recruit between 35 and 40 participants.

6.2.2 *Materials*

6.2.2.1 *Behavioural Task*

The experiment was created using Presentation software. All word stimuli were presented in black type using Monaco font on a plain white background, in the centre of a computer monitor (Intel® Pentium® 4 CPU 2.80GHz, with a 15 inch VGA CRT monitor with standard keyboard attached, which was used to register participant responses). A fixation cross was displayed in the centre of the screen between word presentations.

The adjective stimuli for the presentation and recall phases of the task were taken from a list of 139 positively and negatively valenced human trait adjectives devised by Cili (2012), selected from a list of 555 such adjectives, originally rated in terms of likeability (Anderson, 1968). Cili had 20 students rate the valence of each word on a scale from -3 (*very negative*) to +3 (*very positive*), and used the mean rating for each word as its valence index. A subset of 48 (24 positive and 24 negative) of the most highly valenced words from Cili's list was selected, for the present study (see Appendix H).

Table 6.1

Demographic Information

Demographic	Student sample (N = 40)
Mean age (SD)	20.18 (3.74)
Age range	18 - 40
Gender (% female)	90%
Country of origin:	
UK	85.00%
Europe (other)	2.50%
Asia	5.00%
Other	7.50%
Ethnic origin:	
White	85.00%
Asian	7.50%
Black	2.50%
Mixed	2.50%
Other ethnic group	2.50%

Examples of positively valenced words selected included ‘honest’, ‘intelligent’ and ‘optimistic’. Examples of negatively valenced words include ‘hostile’, ‘selfish’ and ‘deceitful’. The mean valence indices for the two sets of words (positive: 2.19; negative: 2.11) did not differ significantly, $t(46) = 1.20$, $p = .24$. Six emotionally neutral words (such as ‘cabinet’ and ‘elephant’), which did not appear in the list developed by Cili, were used for practice trials prior to the main stimuli presentation phase of the task.

6.2.2.2 *Cognitive Fusion Questionnaire,* (*CFQ*; Gillanders et al., submitted). See Chapter IV, Section 4.2.2.

6.2.2.3. *Brief demographics questionnaire*. Participants were asked to indicate their age, gender, country of origin, and ethnic origin. (Appendix I).

6.2.3 *Procedure*

University of Southampton Psychology ethical approval was obtained for this study prior to recruitment via the psychology student credits webpage.

The task consisted of three phases; an instruction/practice phase, a word presentation phase, and a word recognition phase.

Instruction and Practice Phase

The computer used for the task was situated on a desk in a quiet, well-lit cubicle. Participants read a study information sheet and signed a consent form (see Appendix J). They then completed the CFQ and the demographic questionnaire, on the computer. The researcher then asked participants to read a set of instructions on the monitor screen, as follows.

> This is an experiment investigating predictors of performance on a memory of words task.
>
> Words will be presented on the screen, one at a time, one after the other.
>
> Your task is to read each word, and when you are ready, strike the RETURN key to see the next word. There will be a recognition memory task involving these words at the end of the experiment.
>
> In between words, there will be a cross in the middle of the screen. Please focus on the cross while you are waiting for the next word to appear, rather than letting your attention wander.
>
> Please ask if you have any questions.
>
> When you are ready, press the SPACE BAR to begin some practice trials.

The instructions were followed by six practice trials, with the researcher still in the room. The practice trials used neutral words (‘rainfall’, ‘elephant’, ‘window’, ‘cardigan’, ‘refrigerator’, and ‘cabinet’). Non-personal words were used so that participants were able to experience the format of the study without being given prior warning about the nature of the main task stimuli. As the participant cleared each word from the screen, it was replaced by a fixation cross in the centre of the screen,

with an inter-trial interval of 2000 ms. Following the practice trials the researcher left the room and the participant read the following instructions on the screen:
"The main part of the experiment will be in the same form as the practice trials. When you are ready, please strike the Space Bar to start the actual experimental trials."

Word Presentation Phase

The 16 positively and 16 negatively valenced words of the main trials were presented in a pseudo-random order (no sequence of more than three words of the same valence). After viewing a word for as long as required, participants used the keyboard return key to remove it from the screen. The duration for which each word was on the screen was recorded. After 32 trials, a further set of instructions appeared on the screen, as follows:
"A series of words will now be presented on the screen, one at a time, one after the other.
"You should use the keys marked 'SEEN' and 'NOT SEEN' to indicate whether you saw the word on the screen presented earlier in the experiment or not."

Word Recognition Phase

Eight positively and eight negatively valenced words were selected from the 32 presented in the main part of the task, excluding the words that had been presented in the first and last eight word presentation trials, to avoid primacy and recency effects. These words were interspersed with 16 previously unseen words (half positively and half negatively valenced), from the original list of 48 adjectives. These 32 words were presented in a pseudo-random order, one at a time. A response of 'seen' or 'not seen' to a word cleared the screen. There was an inter-trial interval of 2000ms, with a fixation cross in the centre of the screen. Once the participant had completed the memory trials, a final message appeared on the monitor screen as follows:
"You have now finished the experiment. Please let the researcher know.
"THANK YOU FOR PARTICIPATING"

Before leaving, participants were verbally debriefed and given a study debrief form (Appendix J).

6.2.4 *Analysis Plan*

In Stage One of the analysis, the raw data were inspected, outliers removed, and normative data (including skew and kurtosis indices) were obtained for each of

the study variables. Stage Two used t-tests to compare mean view times for positively and negatively valenced words. Stage Three utilised Pearson's r correlation tests to examine the relationships between CF (score on the CFQ) and word view time (EA). PASW Statistics 19 software was used for all data analysis.

6.3 Results

6.3.1 *Stage One. Data Screening and Preliminary Analyses*

Before conducting any statistical tests, the CFQ and behavioural task data were screened for missing responses. There was no data loss from either the CFQ or the behavioural task variables.

View time and memory data were then prepared for analysis. As is common practice in related research fields such as attentional bias research, view times more than 2 standard deviations outside of each participant's mean were excluded, as were reaction times that were below a cut-off point; in this case, 500ms (e.g. Lees, Mogg & Bradley, 2012; Vassilopoulos, 2005). On these grounds, six participants were excluded from the study because more than 25% of their view time data was below 500ms, suggesting that they were not effectively engaged with the task. This left a final N of 34.

Memory data were collected in the form of hits (H) (correctly indicating that a word had been seen in the main word presentation trials), and false alarms (F) (indicating a word had been seen before when in fact it had not previously been presented) as these are the two types of response that are used to calculate recognition memory sensitivity, based on signal detection theory (SDT; Green & Swets, 1966). As is usual practice (Stanislaw & Todorov, 1999), rates of 1 and 0 were adjusted to allow for the calculation of memory sensitivity values. Rates of 0 were replaced with $0.5/n$, and rates of 1 were replaced with $(n - 0.5)/n$, where n is the number of seen previously or not seen previously trials. An index of recognition memory sensitivity (discriminability index; d') was calculated for each participant by entering their F and H scores in an on-line calculator (Neath, 2012), which used the formula $d' = z(H) - z(F)$, where z(F) is the z score for false alarms and z(H) is the z score for hits (Macmillan, 1993).

Consistent with the criteria used elsewhere in this thesis (see Chapter IV Section 4.3.2), a cut-off point of above or below +/- 7 was used to determine unacceptable levels of skew or kurtosis. Table 6.2 shows that none of the study variables showed problematic distributions. Table 6.3 shows the means and standard deviations for all study variables. The mean score on the CFQ (41.24) is almost identical to that found with other student and non-clinical samples.

Table 6.2

Skew and Kurtosis Data for All Study Variables

Variable	skew	c.r.	kurtosis	c.r.
CFQ	.20	.47	-.06	-.07
Pos VT	1.48	3.53	1.36	1.61
Neg VT	1.63	3.88	2.00	2.49
d' Pos	-.75	-1.78	.67	.79
d' Neg	-.99	-2.35	1.32	1.58

Note. CFQ = Cognitive Fusion Questionnaire; Pos VT = view time for positive words in ms; Neg VT = view time for negative words in ms; d' Pos = d' for positive words; d' Neg = d' for negative words; c.r. = critical ratio

Table 6.3

Means and Standard Deviations for All Study Variables

	Community Internet sample (n = 34)	
Variable	Mean	SD
CFQ	41.24	9.43
Pos VT	2084.18	1441.77
Neg VT	2100.04	1285.77
Pos d'	2.22	.61
Neg d'	2.18	.61

Note. CFQ = Cognitive Fusion Questionnaire; Pos VT = view time for positive words in ms; Neg VT = view time for negative words in ms; Pos d' = d' for positive words; Neg d' = d' for negative words; SD = standard deviation

6.3.2 *Stage Two. Comparisons between view time for positively and negatively valenced words.*

A paired sample t test indicated that there was no significant difference in view time between positively and negatively valenced words (Table 6.4).

Table 6.4

Paired Sample t Tests Comparing Positive and Negative View Time

	Word valency				
	Positive Mean (SD)	Negative Mean (SD)	*t*	*df*	*p*
VT	2084.18 (1441.77)	2100.04 (1285.77)	-.30	33	.76

Note. VT = view time in ms

6.3.3 *Stage Three. Relationship Between CFQ Score and Stimulus View Time.*

Table 6.5 shows the bivariate correlation between CF and view time. The relationships between CF and view time for both positive and negative words were statistically significant, with higher score on the CFQ being associated with longer view times for both types of words. A Hotelling-Williams test using Fisher's transformation of r values to Z scores was used to assess whether there was a significant difference between the correlation between CF and view time for negative words and the correlation between fusion and view time for positive words. The results indicate that the difference was not significant, $z = -1.44$, $p = .15$.

Table 6.5

Bivariate Correlations between CFQ Score and View Time

	Pos VT		Neg VT	
	r	p	r	p
CFQ	.42	.014	.46	.001

Note. CFQ = Cognitive Fusion Questionnaire; Pos VT = view time for positive words; Neg VT = view time for negative words

6.3.4 *Memory Data Analysis*

Given that the memory data from the experiment were collected, the relationship between CFQ score and recognition memory (d') was tested, in relation to positively and negatively valenced words. Bivariate correlations showed a non-significant trend for CF to correlate with recognition memory for negative words (r = .28, p = .11), but there appeared to be no relationship between CF and memory for positive words (r = .13, p = .46). However, stimulus view time is likely to have confounded these results, and the study was not sufficiently powered to statistically partial out its effect.

6.4 Discussion

6.4.1 *Study Findings*

This experiment was designed to pilot a new measure of behavioural avoidance and to use this measure and the CFQ to examine the relationship between CF, and behavioural avoidance in relation to self-referential adjectives. It was predicted that fusion would be associated with avoidance of negatively valenced stimuli, indicated by shorter view times for those words. In fact the opposite was found; there was a highly significant correlation between CFQ score and view time for negative words, with higher fusion being associated with longer view times. Both ACT theory and previous relevant research (e.g. Healy et al., 2008) suggest that CF and EA are positively correlated, so it is unlikely that the results of this study indicate

that CF and the behavioural aspect of EA are negatively correlated. It seems more likely that the computer task functioned as a behavioural measure of CF. That is, when participants were given control over how long they viewed each word, view time did not function as a measure of avoidance or willingness to be in contact with the stimuli, but rather it indicated entanglement with negative stimuli and thus difficulty in disengaging from them.

It was more difficult to predict what relationship if any there would be between CF and avoidance/willingness in relation to positively valenced words, because most ACT theoretical writing focuses on fusion with aversive private experiences. The only relevant published study that has included positively valenced stimuli (Healy et al., (2008), was designed to test the impact of a defusion exercise, which had little effect on relationship to those positive stimuli. Furthermore, due to a design flaw in their measurement of thought believability (used as a proxy for CF) in the defusion condition of their study, conclusions cannot be drawn from their thought believability data, including conclusions about relative believability of positive and negative thoughts.

In the present experiment it was predicted that CF would be positively correlated with willingness to view positively valenced words, as indicated by view times, and indeed a significant relationship of this kind was found between score on the CFQ and positive word view time. However, as discussed above, it seems likely that the task was actually functioning as a behavioural measure of CF rather than avoidance/willingness. Therefore, the findings could possibly be interpreted as indicating that people with a higher propensity to fuse do so with both positively and negatively valenced self-referential words (though further testing of this conclusion is needed). If this is the case, it is not that these findings indicate that fusion is related to willingness to view positive stimuli, but that people who score higher on a measure of CF appear to get more ‘entangled with’ and ‘caught up in’ (to use the wording of CFQ items), verbal stimuli, regardless of the valency of those stimuli.

Interestingly, Cochrane et al. (2007) found that in the ERP aspect of their study, participants scoring highly on the AAQ were possibly engaging in more verbal activity than those with low scores, and this activity did not appear to differ in relation to aversive or neutral stimuli. The authors interpreted this verbal activity as cognitive avoidance strategies, but the ERP data offers no possibility of establishing if this was the case, or if this verbal activity could in fact be characterised as CF.

The relationship between score on the CFQ and view time for negative words was not significantly different to the relationship between CFQ score and view time for positive words, based on the Hotelling-Williams correlation difference test. However, the study was underpowered for this test. If a well-powered study did demonstrate a significant difference in these relationships, with the relationship between CFQ score and view time for negative words being stronger (the very small difference in correlation size in this study might suggest this possibility), this could indicate that negative self-referential words form the basis of a superior behavioural measure of CF, compared to positive words, perhaps because we more readily fuse with negative stimuli, as assumed in the ACT literature. An alternative explanation for such a difference might be that there is some bias towards fusion with negative stimuli in the CFQ items.

The results of this experiment indicated that there was no significant difference in view time for positively and negatively valenced words. This is in keeping with findings in the attentional bias literature, where attentional bias for threat is not usually found with non-clinical samples (Bar-Haim et al., 2007).

Although this experiment yielded some interesting results, the computer task did not appear to be functioning as conceived; that is, as a measure of behavioural avoidance. However, a behavioural measure of CF is a useful addition to ACT research resources, complementing, as it does, the CFQ. The ability to examine a construct via more than one methodology can only increase confidence in that construct (Campbell, & Fiske, 1959), notwithstanding the fact that the development of this measure was not the original aim of the study.

6.4.1.1 *CF, EA, and the ACT Model*

The findings of this study are also important because they raise a significant issue regarding the ACT model, in that they can be interpreted (based on the model) to suggest that the computer task measured CF rather than EA, as originally predicted (also in line with the ACT model). Having both processes in the model, one (CF) that implies attentional entanglement with stimuli and therefore difficulty disengaging from them, and the other (EA) implying the need to disengage from stimuli, means that, as was the case here, the model can account for stimulus-related behaviour, regardless of whether the data indicates approach or withdrawal in relation to that stimulus.

It makes theoretical sense to suggest that CF leads to EA (Greco et al., 2008; Pistorello et al., 2000), implying that CF happens prior to efforts to avoid a stimulus and its psychological impact. If it could be demonstrated that these processes occur at different time points following presentation of a stimulus, this would sharpen the predictive power of the model, in that it could be predicted that at time X CF would occur, whereas at (later) time Y, EA would be expected. This would represent an improvement on the current situation where the model predicts both/either processes.

These processes have not been examined in this way to date within the ACT literature, but a similar issue has been explored empirically in a related field. In the attentional bias literature, the vigilance-avoidance model (Mogg & Bradley, 1998), which states that for anxious people vigilance for threatening stimuli is followed by avoidance of those stimuli, has been tested indirectly by examining sequential points in time on the 'attentional timeline' (e.g. Mogg & Bradley, 2006). By presenting stimuli for shorter and longer lengths of time, and finding that attention is turned towards threat stimuli with shorter presentation times, and turned away from those stimuli at longer times, the researchers infer that avoidance follows vigilance in the seconds following presentation of an aversive stimulus. Mogg and Bradley (1998, p. 837) suggest that anxious individuals may "initially direct their attention to threat, but then try and avoid detailed processing of it in an attempt to minimise their discomfort". The avoidance aspect of this description fits well with definitions of EA such as that of Hayes et al., (1996), quoted in Section 6.1.2. More recently, several studies (e.g. Fox, Russo, Bowles, & Dutton, 2001; Koster, Crombez, Vershuere, & De Houwer, 2004) have yielded findings suggesting that for anxious people it is difficult to disengage from negatively valenced stimuli once they are paying attention to them, and it is this that results in longer reaction times in dot-probe tasks rather than vigilance for threat, as claimed by Mogg and Bradley. From this latter standpoint, the vigilance-avoidance model begins to resemble a CF-EA model, as difficulty to disengage closely resembles being entangled with and dominated by stimuli, phrases that are commonly used to describe CF. For possible ways of taking this line of research forward, see Section 6.4.2.

Although the findings of the present study are limited, what this study has achieved is to pilot the kinds of behavioural paradigms that could be used to test these issues further, under controlled, laboratory-based conditions. These observations also point to a useful distinction between treating CF as a trait (as is the case with the

CFQ), and attempting to directly tap into fusing behaviour, moment by moment. Participants who reported via the CFQ that they tended to fuse with thoughts (CF as a trait), in fact exhibited fusion-like behaviour in this study, but might be predicted to exhibit avoidance-like behaviour at different points in time.

Finally, in terms of study findings, there was a non-significant trend for a correlation between CF and recognition memory for negative words, but it is likely that view time had an impact on this relationship. Due to the sample size, the study was statistically underpowered to partial out the potentially confounding effect of view time. In order to more accurately assess the relationship between CF and recognition memory, a larger sample would need to be recruited for the current experiment, to achieve enough statistical power to be able to control for the impact of view time. Given that the task was shown not to be functioning as originally predicted, it was decided not to continue recruiting. An alternative means of accurately testing the CF-memory relationship would be to experimentally control view time, which as outlined in Section 6.4.2, is likely to be a design feature of any future developments of this experiment.

6.4.2 *Methodological Limitations*

There were a number of limitations, both to the overall design of the study, and specifically to the design and functioning of the behavioural paradigm. Firstly, there are limitations related to the sample. The experiment was underpowered (due to sample size), to carry out some secondary statistical tests. It was based on a predominantly female, student sample, with an under-representation of people from ethnic minorities. These limitations should be borne in mind when generalising from the findings.

With regards to the functioning of the behavioural task, it is commonly the case that when developing a new behavioural paradigm, a phase of testing, modifying and retesting is necessary before the task under development functions as required. This study represents the beginning of such a process, and as such it was expected that further ‘fine-tuning’ of the task design would be required. In the event, as outlined in Section 6.4.1, substantial modifications appear to be warranted. The methodologies used in the attentional bias literature may suggest ways to test the predicted differential impact of CF and EA on behaviour, by developing dot-probe or

other tasks that examine behaviour at several time points following stimulus presentation. The limitation of the dot-probe methodology however is that as the name implies, it measures response time to a neutral dot probe, and infers attentional biases from this data. A strength of the task used in the current experiment was that it yielded a more direct measure of avoidance/approach behaviour. However, it was designed to allow participants to view stimuli for as long as they wished, and therefore view time could not be experimentally manipulated. The task would need to be significantly modified to render it capable of measuring behavioural responses at different time points.

In the current experiment, the mean stimulus view time was approximately 2000ms, and at that point on the attentional timeline (using attentional bias literature terminology), higher score on the CFQ was positively associated with view time for negative stimuli, suggesting psychological entanglement with those stimuli. One possibility would be to conduct a second set of stimulus presentation trials for each participant, with the stimulus presentation time (determined by the experimenter) varying around a fixed time substantially greater than 2000ms, that is, further along the time line. Obviously view time could not be used as a behavioural measure, but the number of attempts to remove stimuli from the screen by hitting a key might be expected to function as a measure of behavioural avoidance at this longer view time.

Given that participants would not actually be able to influence how long the stimuli were presented for, efforts would need to be made to convince them that they did have control over stimulus view time so that they would attempt to remove stimuli from the screen. Having the presentation time vary would be one way of making the design of the experiment less obvious. Another would be to have practice sessions in which they did have control over presentation time, immediately prior to the experimental trials.

It would be hypothesised that at this later point on the time line, score on the CFQ would correlate with attempts to remove negative stimuli from the screen, the latter functioning as a measure of the behavioural aspect of EA. In fact, using this modified version of the behavioural task, latency to first key strike (to remove the stimulus from the screen) and number of key strikes could both be measured, with the prediction that latency to first key strike would function as a measure of CF, and number of subsequent key strikes as a measure of EA. If an orderly change in behaviour in relation to change in stimulus duration could be demonstrated, this

would help address the problem of the ACT model being able to account for both entanglement with and avoidance of aversive stimuli, outlined in Section 6.4.1. The CFQ, which has been shown to be a valid measure of CF, would be essential to test these predictions.

If CF was shown to occur before EA, this says nothing about the former causing the latter. All that could be concluded from this type of study design is that the findings were consistent with the hypothesis that CF leads to EA. An experimental manipulation of CF would be required to demonstrate causality.

6.4.3 *Implications and Future Research*

The results of this experiment represent a unique contribution to ACT research, in that they demonstrate the initial work to develop novel behavioural measures of key ACT variables. The development of easy to administer, validated behavioural measures of CF and EA would significantly broaden the range of methodological possibilities for ACT research. To a large extent ACT-focused research is currently dependent on self-report measures and, in the case of EA, a less than ideal self-report measure. To be able to measure behaviours in relation to CF and EA, rather than rely on verbal self-reflections, would be wholly in keeping with the theoretical and philosophical underpinnings of ACT, and it has long been understood (Campbell & Fiske, 1953) that to have confidence in the validity of a construct and measures of that construct, methodological heterogeneity in measurement design is important. To have available good quality self-report and behavioural measures of core ACT variables would be an important development, and one to which this study makes a contribution.

The findings from this experiment provide preliminary empirical evidence relating to a common assumption in the ACT literature and in ACT therapy (that we can fuse with both positive and negative material, but more readily with the latter). The study findings also highlight important issues for the theory on which ACT is based, and as such contributes an essential step in the scientific testing of ACT as a psychotherapy (David & Montgomery, 2011; Lohr, 2011). Although preliminary in nature, the findings of this study raise interesting questions in that they could be explained in a number of ways.

In addition to the modifications to address design and performance limitations suggested in Section 6.4.2, future research could involve other possible enhancements and developments being made to the behavioural task and to the experiment as a whole. In terms of the task, introducing neutral stimuli in addition to positively and negatively valenced stimuli would allow testing hypotheses to do with the relationship to emotionally charged stimuli compared to neutral stimuli. Different forms of stimuli such as images of scenes and faces should be tested in similar experiments, to see if the results found in the current study apply only to linguistic stimuli, or are more broadly applicable.

Given that CF and EA are observed in psychotherapeutic settings with people with mental health diagnoses, and that these processes appear (based on clinical observation) to contribute to the maintenance of those mental health difficulties, it will be important to carry out these kinds of studies with relevant clinical populations, to enhance understanding of the impact of CF and EA. In particular, fusion with and attempts to avoid negative thoughts and memories are commonly observed with people with PD diagnoses. More specifically, people with these kinds of mental health problems report particular difficulties in relation to fusion with negative or judgemental self-referential words, thoughts and memories, and this fusion appears to understandably give rise to a range of efforts to avoid such private experiences. Some of the avoidance strategies used by people with PDs, such as self-harm and substance misuse, significantly increase risk to their safety. Other EA strategies, such as avoiding leaving one's home, or avoiding social contact, although not risky in the same way, are clearly debilitating and highly detrimental to wellbeing and quality of life.

Refining the behavioural task developed in the current study and then administering it with PD patients would help to clarify the role of these particular variables with this patient group. The findings of such studies might influence ACT therapeutic interventions. Such findings would also provide empirical evidence relating to another aspect of ACT theory; namely that these processes are universal in nature, and that people with a range of psychological difficulties as well as those without any psychiatric diagnosis, are likely to be impacted by these processes in similar ways, albeit to differing degrees. Care would need to be taken when using these kinds of behavioural measures with a PD sample, given that negative self-referential stimuli are used, but some defusion coaching could be included at the end

of the behavioural task to guard against any potential negative impact. It might be predicted that larger effects would be detected in these kinds of tasks when administered to people with PD diagnoses (compared to this non-clinical, student sample, because of their apparent greater propensity to fuse and avoid.

Finally, as noted in Section 6.4.1, there are similarities between attentional bias models that have been developed in relation to anxiety disorders, and the hypothesised relationship between CF and EA, as it is understood within the ACT literature. It may be fruitful to both areas of research to empirically examine these possible similarities. One strategy for this, described in detail in Section 6.4.2, might involve a modified version of the behavioural task, where stimulus avoidance behaviour would be monitored when stimuli were presented for varying presentation times. This would serve to link the examination of fusion with an existing and well-developed body of research. Another possible way of understanding cognitive fusion and the length of time spent attending to verbal stimuli might be in terms of stimulus elaboration. Again, there is a body of literature addressing this phenomenon, which again could guide the development of future laboratory-based empirical examination of fusion. Developing links in this way between ACT research and other well-established theoretical and methodological approaches would guard against ACT theory being developed in isolation, and thus would reduce the risk of 'reinventing the wheel'.

6.4.4. *Summary*

This experiment demonstrated a unique application of the CFQ in the initial development of new behavioural measures of key ACT processes. The study resulted in the development of a behavioural measure of CF. Considerable further development and piloting work would be needed to modify the behavioural task involved to function as both a measure of CF and EA. The results of this experiment suggest a clear strategy for this future research.

Chapter VII

Study 4: Uncontrolled pilot development trial of an ACT-based group intervention for post-DBT patients with poor personality functioning[6]

7.1 Introduction

As reviewed in Chapter I, DBT is currently the psychosocial treatment of choice for PD, although some DBT graduates continue to experience difficulties post-treatment (see Section 7.1.1). Most psychotherapy treatments for PD (including DBT) have been developed for one particular PD diagnosis, although many patients present with several PDs (McGlashan et al., 2000), suggesting general poor personality functioning across diagnostic categories. Taking into consideration the current strengths and limitations of DBT in relation to patients with poor personality functioning and histories of self-harm, a logical next step might be to develop a post-DBT intervention that would be theoretically compatible with DBT, that would be expected to address poor personality functioning across PD diagnoses, and that would be designed to impact engagement in life as well as symptomology.

Chapter V demonstrated that CF, a key ACT process, is implicated in poor personality functioning. Other research has confirmed the relevance of EA (another central ACT process), to PDs (see Section 7.1.2). These processes are hypothesised to be universal and to underpin many, apparently disparate, psychological difficulties. ACT was developed to impact positively on these processes and should therefore have a beneficial effect across mental health diagnoses. Furthermore, there would seem a strong possibility that ACT might be beneficial for people specifically with poor personality functioning across PD diagnostic categories. Supportive of this hypothesis is the fact that ACT emphasises engagement in a personally valued life[7], an issue with which many DBT graduates appear to struggle (see section 7.1.1).

This study was therefore designed to pilot a form of ACT for patients with general poor personality functioning, who have graduated from DBT and are no longer self-harming, but who are still experiencing difficulties.

---

[6] See Chapter I, Section 1.1.1 for a discussion of usage of the term *poor personality functioning*

[7] The term *personally valued life* and other similar phrases are used in the ACT literature to indicate a life that is meaningful and engaging to the individual. See Section 2.2.1.2 for a definition and a description of values as understood in ACT.

7.1.1 *DBT and PD*

As discussed in Chapter I, DBT is an effective intervention for the reduction of parasuicidal behaviours for people with BPD, and results in reductions in Axis I symptomology for the same patient group. There are also some indications that DBT can be beneficial for people with treatment-resistant depression with co-morbid cluster A and C PDs, with a modified form of DBT currently being tested with this patient group (Lynch et al., 2011-2016). Despite these important benefits of DBT, both empirical findings (e.g. McMain et al., 2009) and the author's clinical experience, indicate that many DBT graduates, although no longer self-harming, continue to lead restricted lives, showing little engagement in personally meaningful activities and relationships, and with continued Axis I and Axis II symptomology.

Linehan (1993) initially suggested that what is currently referred to as DBT for BPD is in fact the first of several stages of therapy required for this patient group. She indicated that many graduates of Stage I DBT would be experiencing "quiet desperation" (Dimeff & Linehan, 2001, p. 2), and would need further therapeutic work to build a life sufficiently personally valued to be "a life worth living" (Linehan, 1993, p. 172). She proposed further stages (once high-risk behaviours had been reduced) to address PTSD symptomology, avoidance of emotional experiences, Axis I disorders, and issues of every-day living, such as relationships and occupation. To date, these follow-on stages of DBT have not been empirically tested and published, but Linehan's outline of them has provided some guidance as to what a post-DBT intervention might usefully target.

7.1.2 *ACT Processes and Personality Functioning*

The findings from Chapter V suggest that CF is associated with poor personality functioning in general (as opposed to being relevant to one specific PD diagnosis). Other research has indicted that EA is also highly relevant to PD. For example, Kingston et al. (2010) found that EA mediated the relationship between known risk factors for problem behaviours such as self-harming and substance misuse, and engaging in those behaviours. These kinds of behaviours are strongly associated with several PD diagnoses. In another mediation study, Gratz, Tull and Gunderson (2008) found that EA mediated the relationship between anxiety

sensitivity and BPD. Also in relation to BPD, Rosenthal, Cheavens, Lejuez, and Lynch (2005) showed that thought suppression (an aspect of EA) mediated the relationship between negative affectivity and BPD symptomology. As part of a DBT outcome trial, Berking, Neacsiu, Comtois, and Linehan (2009) found that EA hampered the reduction of depressive symptomology in patients with BPD diagnoses being treated with DBT.

Outcome studies that address mindfulness, another process relevant to ACT, have demonstrated benefits to participants with complex psychological difficulties such as recurrent depression (e.g. Kuyken et al., 2008), and multi-substance misuse in an incarcerated sample (Bowen et al., 2006). Both problems are highly comorbid with PD.

To summarise, although there are few published studies that directly test the relationship between ACT model processes and personality functioning, the evidence available suggests that these processes are relevant to such diagnoses, and as a result, a small number of clinical studies have been conducted, testing the impact of ACT for PD.

7.1.3 *ACT-Based Interventions for PD*

As reviewed in Chapter II, Section 2.2.1.4.1, three small-scale outcome trials have tested an ACT-based intervention for PD (Gratz & Gunderson, 2006; Hurley & Holmes, 2010; Clarke et al., in prep). Two trials focussed on BPD; both yielded promising findings in relation to Axis I symptomology (Gratz & Gunderson, and Hurley & Holmes). The former also reported significant improvements in BPD symptoms. Clarke et al. reported that their ACT group intervention for a "treatment-resistant" sample, half of whom met the diagnostic criteria for at least one PD diagnosis, outperformed TAU on a range of Axis I, Axis II, and quality of life measures.

However, no study to date has tested ACT with a heterogeneous group of PD patients, with poor personality functioning across PD diagnoses. Although the Gratz and Gunderson (2006) study suggests that ACT might have a positive impact on self-harm, the empirical support for DBT reducing parasuicidal behaviours if far more

substantial, and therefore DBT should initially be offered to people with poor personality functioning who are self-harming and/or experiencing suicidal urges.

However, once they have achieved stability in terms of parasuicidal behaviours, an ACT-based intervention might represent a logical next therapeutic step for this patient group. Such an intervention would be offered with the aims of facilitating engagement in a valued life and improvement in psychological functioning, whilst maintaining safety. The aim of this study therefore was to carry out initial development of an ACT-based intervention for people with poor personality functioning, and to test the pilot form of the treatment with a small, heterogeneous, post-DBT PD sample, examining the feasibility, acceptability and impact of the intervention.

It was hypothesised that there would be significant changes, pre- to post-treatment, on measures of engagement in life, psychopathology, and processes hypothesised to mediate such changes.

7.1.4 *Methodological Considerations*

There have been no previous trials of ACT for post-DBT patients with poor personality functioning, so decisions about the study design and the intervention were guided by clinical understanding of this complex and potentially risky patient group, along with data from previous, similar studies. Published guidance suggests that an initial treatment development trial, particularly with a high-risk patient group, should be small-scale (Campbell et al., 2000), and perhaps in the form of an uncontrolled, pre-post trial (Rounsaville et al., 2001). This is a tried and tested DBT treatment development strategy for patients with complex mental health difficulties (e.g. Telch, Agras, & Linehan, 2000; Goldstein, Axelson, Birmaher, & Brent, 2007), and has also been successfully used as a first step in the development of an ACT treatment protocol for treatment resistant mental health problems (Clarke et al., 2012). The present study was therefore designed as an open, pre-post trial.

Treatment development trials are used not only to pilot a psychosocial intervention, but also to evaluate the sensitivity and applicability of outcome and process measures for the patient group. In total, 10 measures were included to evaluate their performance (See Section 7.2.3). ACT targets psychological

inflexibility, in the service of increasing valued action and engagement in life. Measures relevant to these aims, such as the Valued Living Questionnaire (VLQ; Wilson, Sandoz, Kitchens, & Roberts, 2010), and the Flourishing Scale (FS; Diener et al. 2009), were therefore used as primary outcomes. It was also considered important to include measures of inflexibility and related processes.

ACT is not designed primarily to reduce symptomology, although psychiatric symptoms often do reduce following ACT interventions. Measures of Axis I and II symptomology, such as the SCL-90-R (Derogatis, 1994), and the SCID-II (First et al., 1997) are commonly used in PD psychotherapy trials, and were therefore included as secondary outcome measures. The use of these common psychometric measures supports the replication of studies and comparison with findings from different studies. Given that most previous relevant studies have addressed a specific PD diagnosis, these trials have not included a measure of personality functioning across diagnostic categories. Taking into consideration the heterogeneous nature of the patient group selected, the SIPP (Verheul et al., 2008) was included as an outcome measure.

It is expected practice in clinical trials to measure clinician adherence to the study treatment protocol (Rounsaville et al., 2001). However, by definition, in early-stage treatment development trials, the protocol has not been finalised and is open to modification. For this reason, clinician protocol adherence was not assessed.

The previously published ACT studies referred to in section 7.1.3 suggest that an intervention that in some way combines ACT and DBT might be effective for this patient group, and all the relevant trials reviewed in Section 7.1.3 were based on group interventions. In fact, given that the clinicians for the present study had already designed and successfully tested an ACT-based group intervention for people with complex problems including PD (Clarke et al., in prep), it was decided to base the ACT intervention for the current study on the existing protocol, but to make some substantial changes including the addition of some elements of DBT, to take into consideration the specific patient group (see Appendix K for details of the protocol development).

Finally, although Gratz and Gunderson indicated that an ACT/DBT intervention might effectively impact self-harm, it was decided to take a conservative approach to the development of this intervention for this particularly complex patient group, by excluding people who had engaged in parasuicidal behaviour, as defined by

Kreitman (1977)[8], within the previous six months. All participants however, had extensive histories of self-harm.

7.2 Method

7.2.1 *Design*

This study was based on an uncontrolled, pre-post design. The IV was a 20-session ACT-based group intervention, with DVs including measures of engagement in life, valued action, personality functioning, depression, global symptom severity, and self-harm frequency. Process variables were measures of CF, psychological inflexibility, and self-compassion. All measures were administered pre-intervention, post-intervention, and at 6-month follow-up, except the SCID-II measure of PD diagnosis, which was administered pre-intervention and at follow-up.

7.2.2 *Participants*

NHS research ethics committee approval (ref: 10/H0502/5; Appendix L), University of Southampton Psychology ethical approval and research governance approval were obtained prior to participant recruitment. Participants were recruited in February and March 2010 via DBT therapists based at the Intensive Psychological Therapies Service, a specialist, tertiary, PD treatment service within Dorset HealthCare University NHS Foundation Trust (for details of the recruitment procedure, see Section 7.2.4.1). Inclusion and exclusion criteria for the study were as follows:

[8] 1. Nonfatal, intentional self-injurious behaviours resulting in actual tissue damage, illness, or risk of death; or 2. any ingestion of drugs or other substances not prescribed or in excess of prescription with clear intent to cause bodily harm or death".

Inclusion Criteria:

1. 18 years old and above.
2. PD diagnosis at intake.
3. Minimum of 12 months DBT prior to this study.
4. No parasuicidal behaviour during the 6 months prior to the study (see footnote in Section 7.1.4 for definition).
5. Continued significant psychological difficulties.
6. Under the care of a community mental health team (CMHT).

Exclusion Criteria:

1. Under the age of 18.
2. Currently, or in the 6 months prior to this study, engaged in parasuicidal behaviour (see footnote in Section 7.1.4 for definition).
3. Currently meets DSM-IV diagnostic criteria for a psychotic disorder.
4. Currently meets DSM-IV diagnostic criteria for eating disorder, with a body mass index currently below 17.5.
5. Currently meets DSM-IV diagnostic criteria for substance dependence.
6. Learning disability.
7. Other organic disorder that might impair capacity to give informed consent, or to participate.

Clarke et al. (2012) detected effects in an ACT treatment development trial for transdiagnostic mental health problems with a sample size of 10. It was therefore decided to aim for the same sample size in the current study. Figure 7.1 shows the flow of participants through the trial stages. Initially, 10 potential participants were approached about the study, of whom six started the intervention. Table 7.1 shows group demographic information, and baseline data regarding engagement in self-harm, DBT, and current mental health status. All participants were female, with extensive histories of deliberate self-harm, averaging 36 years. The period free of self-harm prior to the ACT group averaged 10 months.[9]

All participants had completed DBT prior to attending the ACT group. The duration of DBT varied considerably between participants (16 to 34 months), as did

[9] Based on conversations that occurred several weeks into the intervention, the therapists thought it likely that participant 4 had self-harmed within the 6 months prior to the group, but had not declared this in order to be included in the intervention. The participant was unwilling to clarify this matter. It was decided to allow her to continue in the study.

the interval since the end of weekly DBT sessions, prior to the start of the ACT group (1 to 77 months). All participants scored in the clinical range for either depression or global symptom severity, or both.

Figure 7.1 *Flowchart of Participant Recruitment, Assessment, and Treatment*

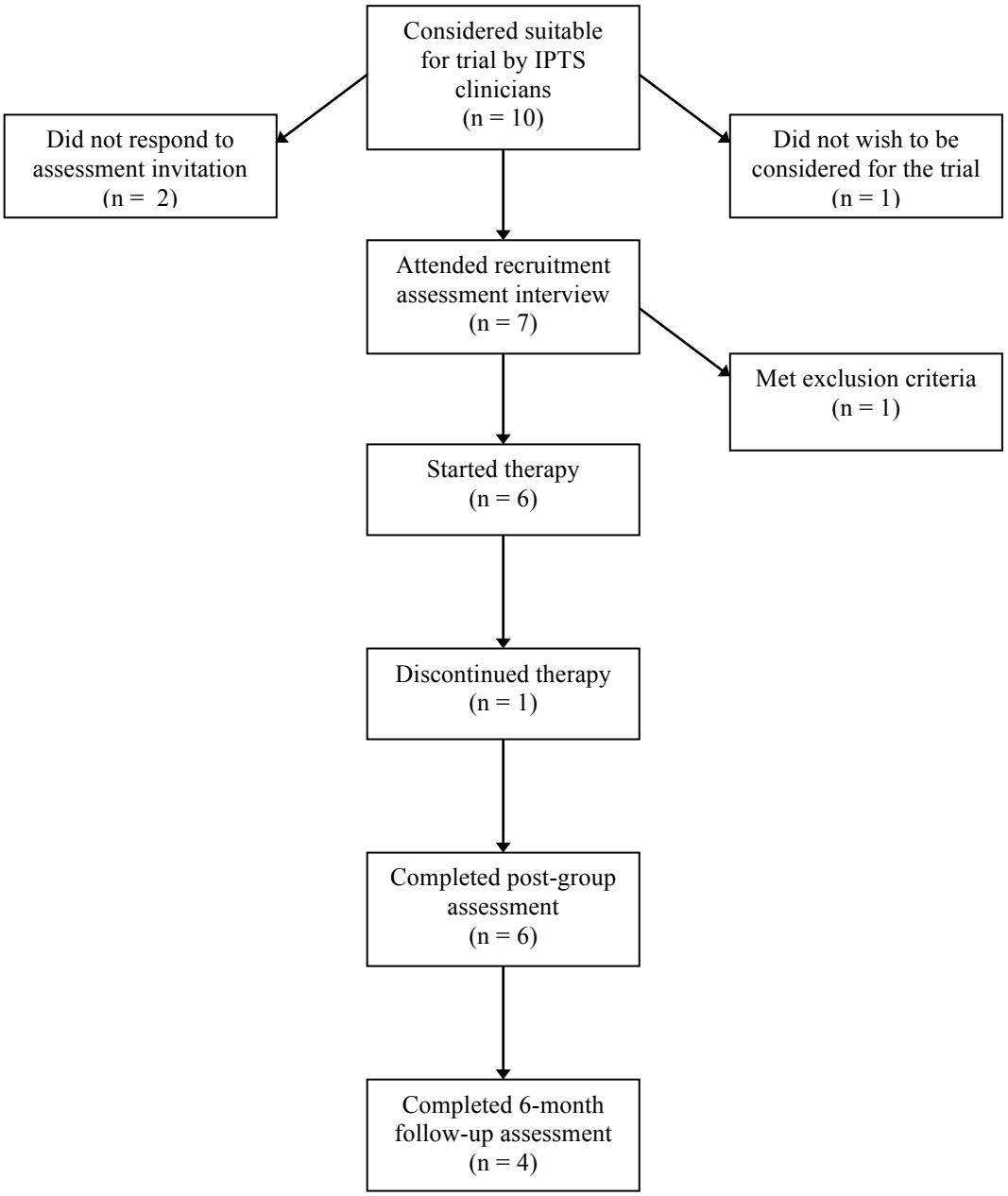


Table 7.1

Demographics and Baseline Statistics

Demographic/Baseline statistic	N = 6
Mean age (range)	46.83 (36 – 57)
Gender (% female)	100%
% taking psychotropic medication	100%
Mean months of DBT (range)[1]	22.40 (16 – 34)
Mean months since end DBT (range)[2]	21.6 (1 – 77)
Mean years self-harm (range)[3]	35.67 (25 – 48)
Mean months since self-harm (range)[4]	10.33 (6 – 14)
% in clinical range for depression[5]	83.33
% in clinical range on SCL-90R GSI[6]	100

Note. [1]Number of months of DBT intervention. [2]Number of months since the end of regular DBT sessions, by the start of ACT group. [3]Number of years history of any form of deliberate self-harm, from first reported instance of self-harm, to most recent. [4]Number of consecutive self-harm free months prior to the start of the ACT group. [5]% scoring at least 14 on the Beck Depression Inventory-II. [6]% scoring at least .70 on the Symptom Check List – 90 Revised, Global Severity Index.

Table 7.2 shows baseline PD diagnosis data for each participant, indicating the very high level of personality psychopathology in the group. All except participant 6 had PD diagnoses from two different PD clusters, with two participants having PD diagnoses from all three clusters.

Table 7.2

PD Baseline Statistics

Participants (N = 6)	PD diagnoses
1	BPD, Paranoid, Dependent, Depressive, Passive-aggressive
2	BPD, Paranoid, Passive-aggressive
3	BPD, Paranoid, Depressive, Passive-aggressive
4	BPD, Paranoid, Schizotypal, Depressive, Passive-aggressive
5	BPD, Paranoid, Dependent, Avoidant, Depressive, Passive-aggressive
6	Dependent, Avoidant, Depressive
Mean PD diagnoses (range)[1]	4.3 (3 - 6)
Group PD diagnoses: (number of participants with diagnosis)[2]	BPD (5), Paranoid (5), Depressive (5), Passive-Aggressive (5), Dependent (3), Avoidant (2), Schizotypal (1)

Note. [1]Lifetime prevalence of specific PD diagnoses based on the Structured Clinical Interview for DSM-IV Axis II Disorders (SCID-II). [2]PD diagnoses in the sample, arranged in order of frequency.

7.2.3 *Materials*

7.2.3.1 *Primary Outcome Measures*

All measures were administered pre-intervention, post-intervention, and at 6-month follow-up, except for the SCID-II PD diagnostic interview, which was administered pre-intervention and at 6-month follow-up only. The SCID-II was administered differently in order to focus the SCID-II assessment on the 12 months prior to each of the assessment dates, that is the year prior to the ACT group starting, and the year prior to the 6-month follow-up date (the latter 12-month period covered the time that the group was running and 6 months post-treatment).

Structured Clinical Interview for DSM-IV Axis II Disorders (SCID-II; First, et al., 1997). The SCID-II is a semi-structured interview designed to diagnose PDs in line with DSM-IV criteria. It consists of a 140-item screening questionnaire, followed by the interview, which is used to examine in more detail positively scored items from the screening questionnaire. Example of items include "Do you find it hard to disagree with people even when you think they are wrong?" (dependent PD), and "Do you have trouble finishing jobs because you spend so much time trying to get things exactly right?" (obsessional-compulsive PD). Items on the screening questionnaire are responded to yes or no. Items in the interview can be rated by the interviewer as 1 – symptom not present, 2 – threshold, or 3 – symptom present. Each PD diagnosis is based on the number of 3 ratings for that PD, with the number required for diagnosis varying from PD to PD. The SCID-II has been shown to have adequate internal consistency (α ranging from .61 to .97 across diagnoses, with a mean α of .82 (Maffei, et al., 1997).

The SCID-II was designed to measure the lifetime prevalence of PDs, as PDs have traditionally been seen as disorders from which it was not possible to recover. In line with previous research (Clarke et al., in prep), the SCID-II has been adapted so that it yields lifetime prevalence diagnoses, as originally intended, but also previous 12 months prevalence data, in order to examine change in PD diagnoses over time.

Severity Indices of Personality Problems (SIPP; Verheul et al., 2008). For details of the SIPP, see Chapter IV, Section 4.2.2.

Valued Living Questionnaire (VLQ; Wilson et al., 2010). The VLQ is the most commonly used values-focused measure in ACT research. It was designed to assess the importance of different life domains to the respondent, as well as how consistently they have acted with respect to each domain in the week prior to completing the questionnaire. The 10 life domains addressed are family relations, marriage/couples/intimate relations, parenting, friendships/social relations, employment, education/training, recreation, spirituality, citizenship/community life, and physical well-being, with importance and consistency being rated from 1 to 10 for each domain. Internal consistency for the importance scale is $\alpha = .77$, and $\alpha = .75$ for the action scale. However, the original VLQ scoring procedure is often modified by researchers and clinicians. For this study, the adaptation developed by Taylor (2010) was used, which focuses on the three life domains rated as personally most important by the respondent.

Flourishing Scale (FS; Diener, et al. 2009). The FS is an 8-item measure of important aspects of human functioning and engagement in life, including meaningful relationships and a sense of purpose. Items include " I lead a purposeful and meaningful life", and "I am engaged and interested in my daily activities". Each item is rated on a 7-point scale, ranging from "strongly agree" to "strongly disagree". There is little published research using the FS to date, but some psychometric data is available. The non-clinical mean score is 44.97 (SD = 6.56). It has good internal consistency (α = .87), and adequate test-retest reliability (*r* = .84), over a 3-month period.

7.2.3.2 *Secondary Outcome and Process Measures*

Symptom Checklist-90-R (*SCL-90-R*; Derogatis, 1994). For details of the SCL-90-R see Chapter IV, Section 4.2.2.

Beck Depression Inventory, 2nd edition (*BDI-II;* Beck, Brown, & Steer, 1996). For details of the BDI-II, see Chapter IV, Section 4.2.2.

Deliberate Self-Harm Inventory (DSHI; Gratz, 2001). The DSHI is a 17-item questionnaire designed to measure the frequency, severity, and duration of several self-harming behaviours. Each item addresses a different form of self-harming behaviour, such as cutting and biting. To measure change in self-harming behaviour over time, Gratz recommends creating a continuous variable based on the number of incidents of self-harm (of any kind), over a specific time period. Based on this variable, Gratz reported a frequency of 1.05 self-harm events per participant over a 3.5-month time period, in a student sample (N = 150). Based on the same variable, the DSHI has acceptable internal consistency (α = .82) and test-retest reliability (*r* = .68 over a 3-week period).

Cognitive Fusion Questionnaire, (CFQ; Gillanders et al., under review) For details of the CFQ, see Chapter IV, Section 4.2.2.

Acceptance and Action Questionnaire Second Version (AAQ-II; Bond, et al., 2011). For details of the AAQII, see Chapter IV, Section 4.2.2.

Self-compassion Scale (SCS; Neff, 2003) The SCS is a 26-item measure consisting of six sub-scales measuring self-kindness, self-judgment, common humanity, isolation, mindfulness, and over-identification. It can also be scored to yield an overall self-compassion score. It is headed "how I typically act towards myself in difficult times", with each item being rated on a 5-point scale, ranging from

"almost never" to "almost always". Examples of items include "I try to be loving towards myself when I'm feeling emotional pain" (self-kindness item), and "When I think about my inadequacies, it tends to make me feel more separate and cut off from the rest of the world" (isolation item). The non-clinical mean for the overall score is 18.96 (SD = 3.64). A mean overall score in an outpatient depressed sample has been reported (Kuyken, et al. (2010) as 14.74 (SD = 5.54). It has excellent internal validity (α = .92) and test-retest reliability (r = .93 over a 3-week period).

7.2.4 *Procedure*

7.2.4.1 *Recruitment*

DBT therapists at the IPTS were asked to identify potential participants from their caseloads. They discussed recruitment with their current DBT patients and nominated interested candidates who met the criteria. They also nominated suitable DBT graduates. Nominated candidates received a patient information sheet (see Appendix M) and an invitation to an individual assessment with a therapist (SC or HB).

These assessments were used to establish whether potential participants met the inclusion criteria or any exclusion criteria, as well as whether the assessor and patient both thought that the intervention would be relevant to the individuals' needs. Potential participants meeting the various criteria were offered a place in the group and written consent was obtained prior to inclusion in the study (see Appendix M). The participant's care network, including GP and CMHT professionals, were sent information about the study (see Appendix M), as well as confirmation that the patient had consented to participate. CMHTs were asked to confirm that they would keep the individual registered to their care for the duration of the study, to ensure a familiar source of support if needed.

7.2.4.2 *Pre-Intervention Assessment*

Immediately prior to the start of the group, participants were sent the pre-group pack of questionnaires to complete, and were invited to attend a pre-group assessment meeting. At the meeting, the SCID-II PD diagnostic interview was carried

out by a suitable qualified and experienced mental health professional from the IPTS team, who was independent to the research team.

7.2.4.4 *Intervention*

The ACT group consisted of 20 weekly, 2.5-hour sessions, and one further 2.5-hour session at 6-month follow up. The general structure of the sessions was as follows:

- Brief mindfulness practice and review of this practice
- Review of diary cards (containing information about risk issues and skills use over the previous week), for each participant
- Review of homework set in the previous session, followed by any teaching points or individual therapeutic interventions arising from this review
- Short break
- Weekly topic (teaching and experiential exercises)
- Mindful review of the session
- Set homework for the following week

The group was facilitated by two ACT and DBT trained clinical psychologists, Sue Clarke (the CI for the study), who is a UK DBT trainer, and Helen Bolderston (PhD student). They provided each other with peer supervision, as well as receiving a small amount of phone supervision from Dr Kelly Wilson, one of the developers of ACT. The group was run by two therapists rather than one, due to the complex and challenging nature of the patient group. This strategy had been effective in the Clarke et al. trial, and is consistent with DBT skills groups. The stated aims of the group were to:

(i) Maintain behavioural stability (in terms of parasuicidal behaviour)

(ii) Begin to develop a life that is more valued and less restricted.

A description of ACT can be found in Chapter II. An outline of the specific topics addressed in each session of this protocol can be seen in Table 7.3. A more detailed description of the treatment phases can be found in Appendix N. The therapists applied the protocol flexibly, making minor modifications depending on their in-the-moment experiences of individual participants and group dynamics. Similarly, although the 20 sessions were designed to guide participants through a

series of treatment phases in a linear fashion, sensitivity to the needs of the group occasionally indicated that topics should be held over until the next session, or revisited more than once throughout the life of the group. Over the course of the intervention as a whole however, all planned topics were addressed.

7.2.4.5 *Role of DBT in the Intervention*

Although the intervention was predominantly ACT-based, some DBT-focussed content and structure was included to help maintain behavioural stability and to facilitate the transition from DBT. The basic sessional structure (see Section 7.2.4.4) resembled a DBT skills-group. Participants completed a DBT diary card each week, which the therapists reviewed with them to monitor risk and DBT skills use. Participants were consistently coached to use DBT skills in challenging situations. ACT was thus introduced as a compatible addition to DBT rather than a replacement, enabling patients to move towards engagement in a valued life, whilst remaining safe.

7.2.4.6 *Modification of ACT*

ACT has a reputation for being a psychologically evocative therapy, in that by encouraging patients to be psychologically present with their private experiences, it almost inevitably results in patients at times experiencing uncomfortable emotions, memories, and so on. Patients in the ACT condition of Clarke et al. (in prep.), for example, reported this aspect of ACT to be both challenging and rewarding. However, people with poor personality functioning tend to be particularly unskilled in terms of emotion regulation and processing, and therefore can become very anxious at even the thought of experiencing emotions. Linehan referred to people with BPD as “emotionally phobic” (Linehan, 1993), for example. When developing the treatment protocol, it was thought that the classic ACT message that control and avoidance of private experiences is problematic and acceptance is the most effective way forward (Hayes & Strosahl, 2004), could be particularly disturbing for this patient group. The anxiety generated might result in patients becoming less psychologically flexible and therefore less willing to engage in valued behavioural change. Such fear of experiencing emotions and other private experiences might also increase the risk of self-harm, which all group participants had used historically to cope with difficult emotions, memories and so on.

For these reasons, this key ACT message was offered in a specific form; that *over-reliance* on avoidance and control strategies was likely to cause long-term difficulties such as lack of engagement in valued living, but that such strategies might be effective in the short-term as survival strategies. The overall aim in relation to this fundamental ACT principle was for patients to develop the ability to discriminate between situations where temporary avoidance of private experiences might be the most effective strategy, and situations where acceptance of such experiences might be most effective, and then being able to act accordingly. This message was designed to be compatible with DBT, where patients are encouraged to use both acceptance and change strategies, with distraction, for example, (a clear avoidance strategy), being viewed as an effective distress tolerance skill. At the same time, the message was seen as being entirely compatible with the ACT principle of workability, and certainly more in line with ACT principles than a simplistic, rigidly held rule to the effect of 'you should never avoid or control your private experiences, regardless of the context in which they are arising.'

Some ACT experiential exercises were also excluded or modified to make them more suitable for this patient group. For example, the values clarification exercise 'Attending your own funeral', which is included in several ACT textbooks and many effective ACT protocols (Hayes & Strosahl, 2004; Hayes & Smith, 2005), was not included, as it was thought that its focus on the imagined recent death of the participant, would be emotionally dysregulating for this patient group. This (and other ACT exercises) were changed not as an avoidance of material that might trigger thoughts of death or suicide (such topics of course came up from time to time in a group for post-DBT PD patients in any case), but so that group participants could engage in the experiential exercises and hopefully benefit from them, rather than merely becoming dysregulated.

It is not just ACT experiential exercises that can be emotionally stirring. The behaviour of ACT therapists can also contribute to patients experiencing intense emotions. The therapist being psychologically present to their own thoughts, emotions and other private experiences during therapy, at times using self-disclosure about those experiences, communicating care and compassion for the patient, being willing to show vulnerability, being willing to notice and stay with pain and discomfort; these valued aspects of ACT therapist behaviour can all serve to intensify the emotional and therapeutic experience (supporting processes such as willingness

and defusion), and are seen as important contributors to therapeutic change (Wilson & Dufrene, 2008). However, for people with long histories of interpersonal and intimacy problems, people who are also emotionally phobic and easily dysregulated, experiencing the therapist in this way could easily feel threatening and emotionally overwhelming.

The therapists in this trial therefore modified their behaviour, in that they used less self-disclosure in relation to private experiences, and disclosed less of their observations or curiosity about patients' in-the-moment psychological experiences. The therapists also conducted fewer individual pieces of therapeutic work in the group (a regular feature of the original 16-week protocol), and those pieces of individual work that did happen were shorter and less exposing for the patient.

7.2.4.7 *Role of Mindfulness in the Intervention*

Each group session began with a brief mindfulness exercise followed by a review of participants' experiences of that exercise. The reason for including these exercises were:

1. To help participants become more psychologically present at the start of the session.
2. To support participants to hold any distress in mindful awareness.
2. To communicate the need for regular mindfulness practice to the group.
3. To keep a familiar feature from DBT.

Mindfulness practices were gradually changed over the course of the intervention to allow participants to build their capacities under relatively undemanding conditions. Thus exercises (a) increased in duration (from 2 – 10 minutes), and (b) moved from a focus on objects or experiences outside of the body (e.g., pebbles) to a focus on internal experiences (e.g., thoughts, emotions, bodily sensations) as group members' skills developed. These transitions were designed to overcome common difficulties in achieving defused, mindful awareness of internal, personal experiences (compared to external, less personal experiences). Significantly, many people with poor personality functioning have experienced body-related trauma and may initially be disturbed by a focus on physical sensations. All participants were given a compact disc containing recordings of the key group mindfulness and defusion practices. These were recorded by HB, an experienced mindfulness teacher.

Table 7.3
Weekly Session Topics

Treatment phase 1: Transition from DBT to ACT

Session 1: A foundation to build on
Topics covered: Introductions; ground rules; orientation to the intervention structure and style; ACT and DBT – similarities and differences; the importance of DBT as a stable base from which to begin engaging more flexibly in life.
Experiential exercise: Pairs-based discussion of 'One thing that I changed through DBT that I am pleased about' and 'One positive step I would like to take in my life during this group'.

Session 2: A brief introduction to values
Topics covered: What values are and are not; compass metaphor; values and goals; values and vulnerability; connecting with your values can be uncomfortable; Why it might be effective in life to connect with values.
Experiential exercise: What matter enough to me that I'm putting myself through 20 weeks of therapy, with all the struggles that is likely to entail?

Treatment phase 2: Creative hopelessness

Session 3: Why language leads to suffering
Topics covered: How difficult it is for humans to be happy; simple, RFT-based explanation of the role of language and cognition in human suffering; language as a gift and a burden; introduction to EA and the potential difficulties it can lead to.
Experiential exercise: Thinking/writing about the kinds of actions they take to supress, control, avoid uncomfortable private experiences, followed by sharing in the group.

Session 4: The pull of avoidance
Topics covered: EA, the pain of presence and the pain of absence; different rules inside and outside the body for control and avoidance; short-term use of avoidance and control strategies for crisis survival vs. costs of persistent/long-term use; why we do what doesn't work.
Experiential exercise: 'Yellow-jeep' thought suppression exercise.

Session 5: Why willingness?
Topics covered: What willingness is and is not; willingness as an alternative to control and avoidance; Chinese handcuffs metaphor; research evidence on the role of EA and willingness.
Experiential exercise: Being willingly out of breath; 'Why and how might willingness be relevant to me personally?' written exercise.

Treatment phase 3: Cognitive defusion

Session 6: The trouble with thoughts
Topics covered: Noticing the process of thinking; CF/defusion.
Experiential exercise: Brief guided mindfulness of thoughts exercise; 'Watching the mind-train' defusion exercise.

Session 7: Having a thought vs. buying into a thought
Topics covered: CF and defusion; the conditioned nature of thoughts; the goal of defusion.
Experiential exercise: Defusion exercise – 'Milk, milk, milk'; Complete the following sentences – 'Mary had a little . . .'; Brief cognitive defusion exercise: 'I am having the thought that . . .'; Cognitive defusion exercise – 'Leaves on a stream'.

Session 8: Review session
Topics covered: No new topic introduced. Participants encouraged to suggest topics they would find it helpful to revisit.
Experiential exercise: None planned – organised in the moment in response to participants' needs.

Treatment phase 4: Developing a different relationship to history, private experiences and self

Session 9: Willingness revisited
Topics covered: Why willingness matters – the possibility of a different relationship with your private experiences and history; willingness and traumatic history; willingness as a path to emotional freedom
Experiential exercise: The unwelcome guest at the party metaphor

Session 10: Mindfulness
Topics covered: Reviewing what mindfulness is; why mindfulness is important; why mindfulness is so challenging; the aims of mindfulness; tips on how to practice.
Experiential exercise: 3-minute breathing space

Session 11: If I'm not my thoughts, then who am I? Self-as-context
Topics covered: Simple RFT explanation of how our sense of self develops in childhood, and what happens when we grow up in less than optimal conditions for this; metaphor of a photographer lighting a scene in different ways so that different aspects are visible and invisible, to explore taking different perspectives on ourselves – some features can become 'highlighted' and we lose sight of others; getting overly attached to one view of the self vs. holding our self-stories lightly; self-as-content; self as a process of on-going awareness; the observing self
Experiential exercise: Exploration of the chess metaphor with a chessboard and pieces; mountain meditation – metaphor of the self as a solid, unchanging mountain in the presence of emotional storms and weather

Treatment phase 5: Values and committed action

Session 12: Reintroducing values
Topics covered: Sweet moments; why pay attention to sweet moments?; blue beads in a jar metaphor; why choose valued living? – research evidence and personal reasons (giving dignity to struggle and pain, building a more satisfying life etc).
Experiential exercise: Pairs exercise describing a recent sweet moment

Session 13: Choosing to value
Topics covered: Values as chosen life-directions – 'passengers on the bus' metaphor; values questions: 'if you could choose, what would you choose?', 'how do you want to be in your life?', 'what do you want to stand for?'
Experiential exercise: 80th birthday values exercise; clarifying values worksheet.

Session 14 What I do and how I feel
Topics covered: The link between activity and mood; low mood and low activity; different functions of activity – satisfaction, pleasure, self-nurturing, depleting, avoiding/numbing.
Experiential exercise: Written exercise noting the function of activities yesterday – discussion in pairs – 'what have I learned from this exercise about how I spend my time and the function of these activities (or lack of them)?'

Session 15 Increasing values-consistent activity
Topics covered: Valued actions – choice and flexible responding vs. rigid rules; values as a process rather than an outcome; barriers to taking valued action.
Experiential exercise: 10 steps to 'trying on a value' – choosing a value to work with, planning to take a small step in that valued direction, dealing with barriers etc.

Session 16: Increasing values-consistent activity 2: Starting to do things you've been avoiding
Topics covered: The metaphor of 'a jump is a jump' – taking valued action is not about the size of the jump, it's about jumping into life, regardless of the size of the jumps; why we don't jump into life – what holds us back; metaphor of the willingness dial; metaphor of the tin-can monster.
Experiential exercise: Planning small steps in a valued direction – something you value but have been avoiding.

Session 17: Review session: The ACT story so far
Topics covered: No new topic introduced. A review of the journey of the group so far; participants encouraged to suggest topics that they would find it helpful to revisit.
Experiential exercise: None planned – organised in the moment in response to participants' needs.

Treatment phase 6: Looking beyond the group

Session 18 DBT and ACT; Acceptance and change
Topics covered: Acceptance and change in both DBT and ACT; 'in any moment, how do I make a choice between acceptance and change?' – current psychological and physical resources, long-term costs vs. short-term benefits; what does my experience (rather than my mind) tell me; chronic or discreet circumstances.
Experiential exercise: Building the habit of taking valued action: 10 steps to 'trying on a value'.

Session 19 The choice to live a more vital life
Topics covered: Metaphor of the crucial fork in the road – the old path, where avoidance is in charge of your life, and a new path where valued living plays a role.
Experiential exercise: Building the habit of taking valued action: 10 steps to 'trying on a value'.

Session 20 The choice to live a more vital life continued
Topics covered: The group ending; reflection on the group experience; future plans
Experiential exercise: Writing exercise focused on – what I most value in life; what I've been doing to cope with difficult thoughts etc. in relation to this, the cost to me and others of me rigidly relying on these coping strategies, what it would mean to be able to respond to life more flexibly, to make choices based on my values and goals, what would a small step in that direction look like and what would it mean to me?

7.2.4.8 *Post-Intervention Assessment*

Immediately post-intervention, patients completed a questionnaire pack. They were also invited to attend an informal meeting with a research assistant working for the IPTS, to gather feedback about their experience of the group, as well as any suggestions for changes they would want made to the protocol for future groups.

7.2.4.9 *Follow-Up Assessment*

A final group session occurred 6 months after the final weekly session. This was used to support participants to re-orient to the aims and methods of the intervention, and to guide them to use ACT and DBT approaches as appropriate, to

address any difficulties they were experiencing. A final set of questionnaires was also administered at the 6-month follow-up, as was a final SCID-II interview.

7.2.5 *Analysis Plan*

Stage One involved dealing with missing data and calculating baseline values for all variables. Group session attendance was also examined to shed light on the feasibility of the intervention. Stage Two consisted of non-parametric group analyses to test for mean differences across the three data collection times, for the primary outcome measures. A non-parametric test was used (Friedman's ANOVA for repeated measures) because of the small number of participants involved. The use of this statistical test, which does not require normally distributed data, meant that the distribution of the variables did not need to be examined. Effect sizes were not calculated due to the small sample size.

Stage Three involved assessing individual participant changes by calculating a reliable change index (RCI; see Section 3.3.1 for an outline of this method). Stage Four involved dealing with data from variables where RCIs could not be calculated due to lack of the necessary psychometric information. Informal, verbal feedback from participants at the end of the intervention was also considered at this stage.

7.3 Results

7.3.1 *Stage One. Preliminary Analysis and Participant Characteristics*

The amount of data missing from the completed psychometric measures was negligible (less than 1%) so missing items were replaced with the sample mean, as recommended by Tabachnick and Fidell (2001).

Patients attended 14 out of 20 treatment sessions, on average. This figure rose to 17 for the five treatment completers. One patient dropped out of therapy after week 7. She completed the post-group, but not follow-up, psychometrics.

Participant 4, in addition to the lack of clarity about her self-harming behaviour prior to the ACT group (see Section 7.2.2), showed an inconsistency in her baseline psychometric data. Her SCID PD assessment (conducted face-to-face) indicated a complex presentation of 5 PD diagnoses, an assessment that matched the

clinical observations of the group therapists. However, at baseline this participant scored just 11 on the BDI-II, a figure indicating no or minimal depression, despite reporting self-harm urges, suicidal ideation, and high levels of misery at times during group sessions, including in early sessions. On other outcome measures at baseline (the SCL90-R GSI for example), she scored in the clinical range, as would be expected from her face-to-face presentation. Her data has been included in the various analyses outlined below, but the reliability of her self-report data in particular remains uncertain.

Table 7.4 shows group mean scores for all study measures at the three data collection time points. Two sets of data were missing from the 6-month follow-up means, due to one participant having dropped out of the study, and another being physically ill and unable to provide data.

7.3.2 *Stage Two. Group statistical analyses*

Friedman's ANOVAs were conducted on all primary outcome variables except the SCID-II. As there were only two data collection points for the SCID-II, a Wilcoxon signed-rank test was used. All participants completing the intervention were included in the analyses (n = 5), with the post-intervention data from participant 2 being carried forward to the follow-up data point, (following recommendations by Hollis and Campbell, 1999), as she completed the intervention sessions but was not able to provide follow-up psychometric data.

As can be seen from Table 7.5, there were no significant group changes in primary outcome measures across the three data collection points. A Wilcoxon signed-rank test indicated that there was also no significant difference between numbers of PD diagnoses based on the SCID-II interview, from pre to 6-month follow-up, ($z = -.54$, $p = .59$)

Table 7.4

Means and Standard Deviations for All Study Measures

Measures	Mean (SD)		
	Pre-group N = 6	Post-group N = 6	6-month follow-up N = 4
Primary			
FS	30.67 (9.22)	35.00 (14.19)	34.33 (12.10)
VLQ CI	6.83 (1.22)	6.83 (1.09)	7.58 (1.26)
VLQ CA	4.47 (2.05)	5.89 (1.63)	6.24 (1.46)
SCID-II	2.83 (2.23)	-	3.25(1.71)
SIPP Self-con	35.00 (7.85)	34.67 (9.14)	29.50 (7.05)
SIPP Respon	37.50 (5.43)	37.67 (6.02)	36.50 (5.80)
SIPP Ident	24.67 (3.83)	28.67 (5.01)	23.25 (.96)
SIPP Relat	27.67 (10.52)	29.67 (11.69)	32.50 (8.96)
SIPP Social	39.33 (5.96)	39.83 (6.79)	35.00 (8.12)
Secondary/ process			
BDI-II	26.83 (10.32)	28.33 (10.73)	34.75 (3.95)
SCL-90R GSI	1.50 (.62)	1.86 (.97)	2.27 (.78)
DSHI	0 .00 (0.00)	1.40 (2.61)	2.80 (4.38)
CFQ	57.17 (18.54)	51.33 (15.02)	56.25 (6.80)
AAQII	35.67 (7.66)	32.17 (9.37)	38.50 (5.97)
SCS	15.71 (3.76)	15.67 (4.21)	13.24 (3.40)

Note. FS = Flourishing Scale; VLQ CI = Valued Living Questionnaire Current Importance; VLQ CA = Valued Living Questionnaire Current Action; SCID-II = Structured Clinical Interview for DSM-IV Axis II Disorders; SIPP Self-con = self control domain; SIPP Social = social concordance domain; SIPP Ident = identity integration domain; SIPP Relat = relational functioning domain; SIPP Respon = responsibility domain; BDI-II = Beck Depression Inventory; SCL-90 GSI – Symptom Checklist-90 Global Severity Index; DSHI = Deliberate Self-Harm Inventory; CFQ = Cognitive Fusion Questionnaire; AAQII – Acceptance and Action Questionnaire 2nd Version; SCS = Self Compassion Scale.

Table 7.5

Friedman ANOVAs for Repeated Measures at Pre, Post, and Follow-Up

	n = 5		
Measures	*chi-sq*	*df*	*p*
FS	.55	2	.76
VLQ CI	1.73	2	.42
VLQ CA	2.00	2	.37
SIPP Self-con	.74	2	.69
SIPP Respon	1.53	2	.47
SIPP Ident	3.11	2	.21
SIPP Relat	.11	2	.95
SIPP Social	4.78	2	.09

Note. FS = Flourishing Scale; VLQ CI = Valued Living Questionnaire Current Importance; VLQ CA = Valued Living Questionnaire Current Action; SIPP Self-con = self control domain; SIPP Social = social concordance domain; SIPP Ident – identity integration domain; SIPP Relat = relational functioning domain; SIPP Respon = responsibility domain.

7.3.3 *Stage Three. Reliable and Clinically Significant Individual Changes*

RCIs were calculated for the following measures: BDI, SCL-90 GSI, SIPP, FS, CFQ, SCS, and the AAQ-II, from pre to post intervention, and from pre to follow-up. RCIs were not calculated for the SCID-II, DSHI, or VLQ, as the required psychometric data was not available for the form of these measures used in the study. The results are displayed in Table 7.6, with those of the primary outcome measures being displayed graphically in Figure 7.2.

An RCI of 1.96 or above indicated a statistically reliable change at the p = .05 level. The clinical change categories were determined (following Jacobson and Truax, 1991) by calculating the mid-point between the clinical mean and the non-clinical mean for each measure, and examining whether each participant's scores had crossed that cut-off point either post-intervention or at follow-up. A reliable change that crossed the cut-off point from the clinical to non-clinical side was labelled

"recovered". A reliable change in a positive direction but not crossing the cut-off point was labelled "improved". A reliable change in the direction of poorer functioning was viewed as deterioration. Non-significant changes were labelled "same".

As can be seen from Table 7.7, the pattern of results is mixed, with some participants improving or recovering on some measures, some deteriorating, and some not showing reliable change. Participant 1 had mixed results immediately post-intervention, including some improvements (the SIPP identity integration subscale and the FS), but appeared to have deteriorated during the follow-up period. Participant 2 showed reliable, positive changes on several measures immediately post intervention, including on the BDI-II, some SIPP subscales, but was not available to provide psychometric data at follow-up. Participant 4 appears to have deteriorated on several outcome measures at the post-intervention point, including the BDI-II and the SCL-90-R, changes that for the most part were maintained at follow-up. Participant 5 showed little change immediately post-intervention, but had improved on several measures by the follow-up data collection point, including three of the five SIPP subscales. Participant 6 showed significant improvement on two primary outcome measures post-intervention, this improvement being maintained at follow-up for the BDI-II only. There were very few changes in score for any of the questionnaires used to address potentially relevant processes (the CFQ, AAQII, and the SCS).

Table 7.6
Reliable and Clinically Significant Individual Changes

	Post (n = 5)				Follow-up (n = 4)			
Measures	Recovered	Improved	Same	Deteriorated	Recovered	Improved	Same	Deteriorated
FS	1	0	4	0	0	0	4	0
SIPP S-con	1	0	2	2	1	0	0	2
SIPP Social	0	0	5	0	0	0	4	0
SIPP Ident	1	1	3	0	0	0	4	0
SIPP Relat	0	0	5	0	0	2	2	0
SIPP Respon	1	0	3	1	0	0	4	0
BDI-II	1	1	2	1	0	1	1	2
SCL-90 GSI	0	0	4	1	0	0	2	2
CFQ	0	0	5	0	0	1	3	0
AAQII	0	0	5	0	0	0	4	0
SCS	1	0	3	1	0	0	3	1

Note. FS = Flourishing Scale; SIPP S-con = self control domain; SIPP Social = social concordance domain; SIPP Ident – identity integration domain; SIPP Relat = relational functioning domain; SIPP Respon = responsibility domain; BDI-II = Beck Depression Inventory; SCL-90 GSI – Symptom Checklist-90 Global Severity Index; CFQ = Cognitive Fusion Questionnaire; AAQII – Acceptance and Action Questionnaire 2nd Version; SCS = Self Compassion Scale.

Figure 7.2

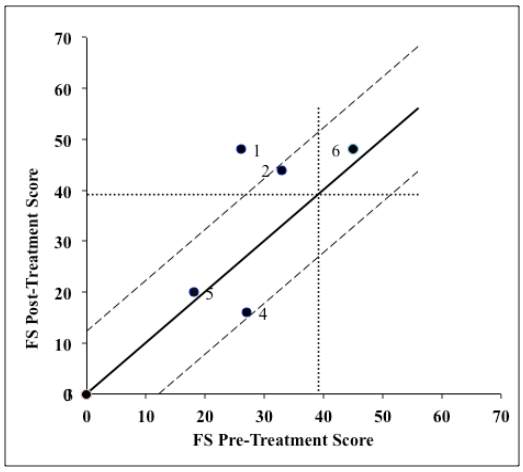
RCI Data for Primary Outcome Measures

Legend

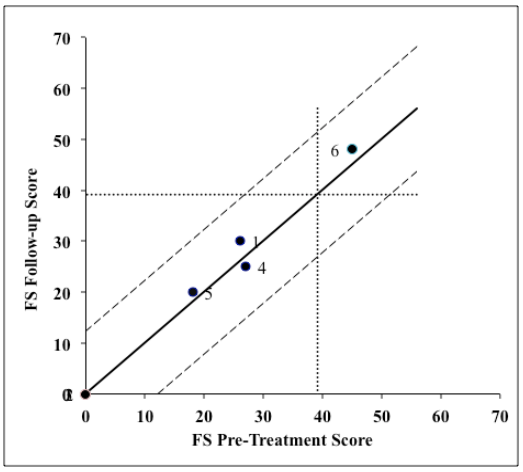
Line of no change Reliable change limits Clinical significance change limits

Flourishing Scale

Pre-Post

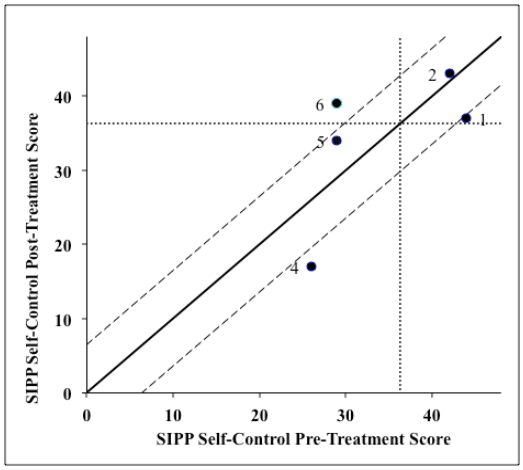


Pre-Follow-Up

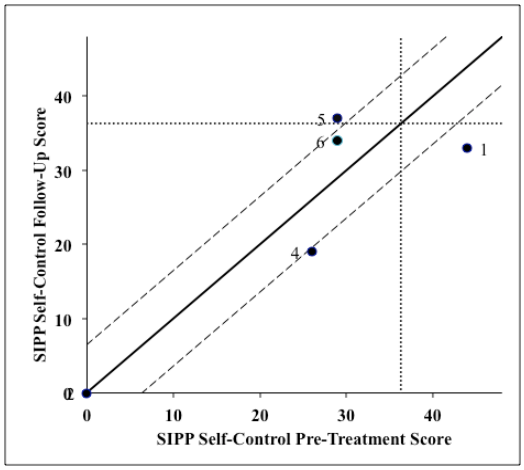


SIPP Self-Control Subscale

Pre-Post

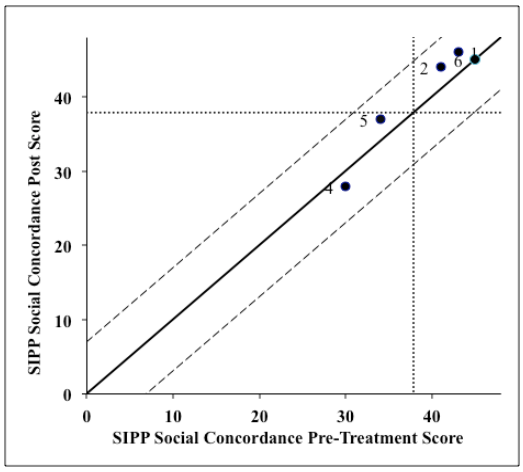


Pre-Follow-Up

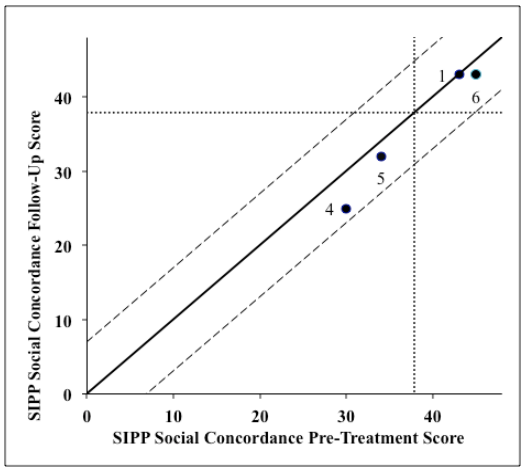


SIPP Social Concordance Subscale

Pre-Post

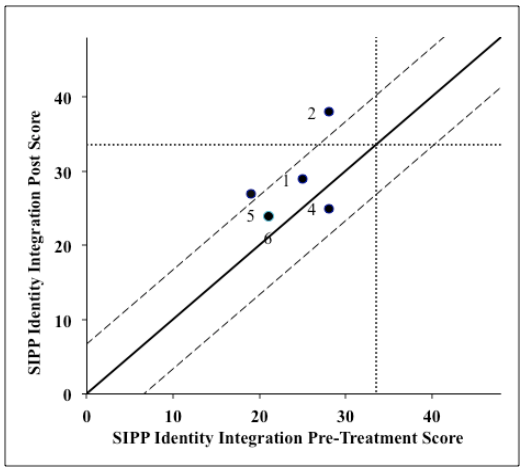


Pre-Follow-Up

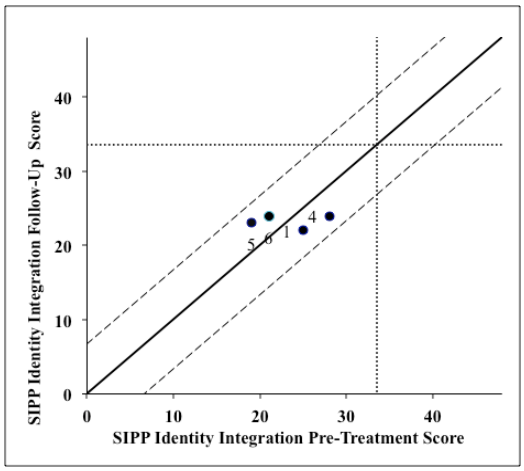


SIPP Identity Integration Subscale

Pre-Post

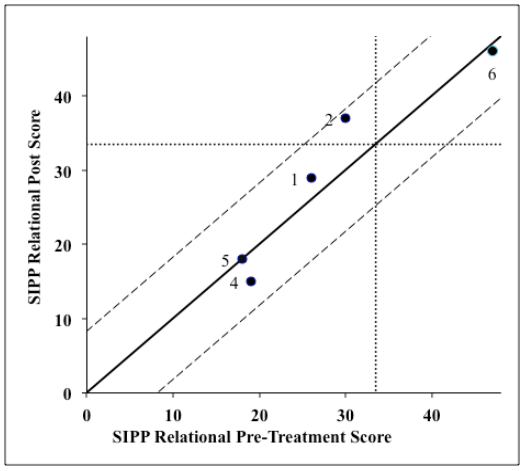


Pre-Follow-Up

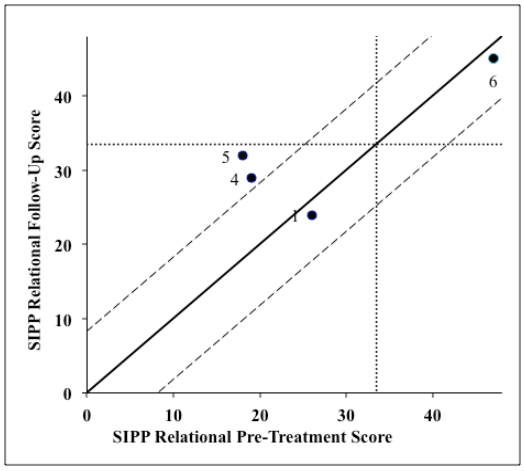


SIPP Relational Functioning Subscale

Pre-Post

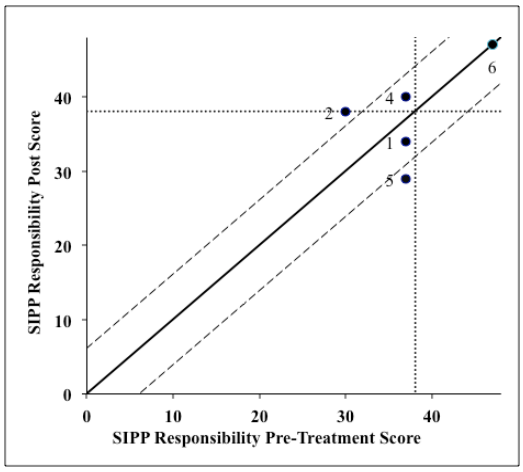


Pre-Follow-Up

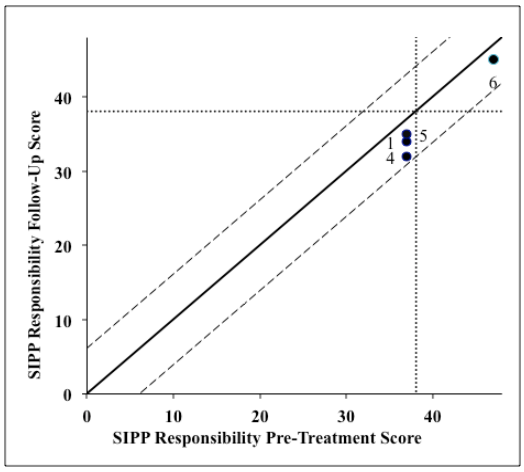


SIPP Responsibility Subscale

Pre-Post



Pre-Follow-Up



7.3.4 *Stage Four. Data Not Analysed Using the RCI Methodology*

The VLQ, SCID-II and the DSHI could not be analysed using the RCI methodology, owing to the lack of required psychometric data for these measures. Figure 7.3 shows group mean changes in current importance of life domains (valuing) and valued action taken, across the three data collection points. The extent to which the group reported valuing the various VLQ life domains remained constant during the intervention, but rose during the follow-up period. Action taken in relation to valued life domains increased during the intervention, and in the follow-up period.

Changes in number of PD diagnoses as determined by the SCID-II are displayed in Table 7.7. As with the RCI data presented in 7.3.3, there was a mixed picture of results, with Participant 1 deteriorating, and Participants 5 and 6 improving. Interestingly, Participant 4, who appeared to deteriorate significantly on several other outcome measures, (though with some doubt about the reliability of some of her self-report data), showed no change in PD diagnoses as determined by this face-to-face interview.

Figure 7.3

Group Mean Current Importance (Valuing) of Life Domains, and Valued Action

Legend

Valued action Current Importance (valuing)

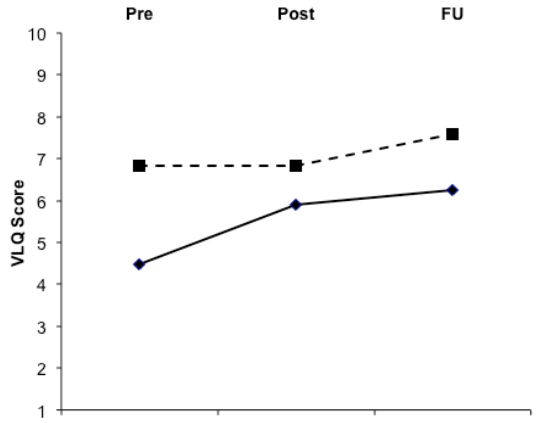


Table 7.7

Changes in PD Diagnoses Based on 12-Month SCID-II Prevalence

Participant	Pre	6-month follow-up
1	2	4
2	0	-
3	2	-
4	5	5
5	6	3
6	2	1
Mean no. PDs	2.83	3.25

Changes in participant self-harming behaviour are outlined in Table 7.8. No participants made suicide attempts either during the intervention or the follow-up period. Participants 2, 5, and 6 engaged in no self-harming behaviour either during

the group or the 6-month follow-up. Participant 4, who is likely to have self-harmed during the 6-months prior to the intervention commencing, also reported having occasionally self-harmed during the life of the group and during the follow-up period. Participant 1 had one minor episode of self-harm whilst the group was on-going, and several more episodes during the follow-up period. None of the participants required medical attention as a result of these self-harm incidents, and using the criteria drawn up for an on-going RCT involving participants with PD diagnoses (Lynch, 2011 - 2016), all incidents of self-harm in the current trial would be categorised as either adverse events or adverse reactions, the two least serious categories of incidents.

Table 7.8

Participant Self-Harming Behaviour

Participant	During therapy	During follow-up
1	1 x scratch	10 x scratch
2	0	0*
4	3 x scratch; 3 x bite	4 x scratch
5	0	0
6	0	0

Note. *Although no psychometric data on self-harm were available for Participant 2 during the follow-up period, her electronic health records were accessed to assess whether any self-harm occurred.

A research assistant carried out informal, face-to-face interviews with each participant, at the end of the intervention, to give them the opportunity to provide feedback about their experience of the group, and any suggestions of changes to be made to the protocol. Feedback was passed on to the therapists in anonymous form. These arrangements were designed to make it easier for participants to give less positive feedback, if necessary. The feedback was not formally analysed, but the most common themes were:

- The intervention was helpful, with mindfulness practice and specific ACT exercises such as cognitive defusion and values-focused behavioural activation, being singled out as particularly useful.
- Although being in a group with other people was anxiety provoking, it was also helpful to get to know other people dealing with similar issues.
- The therapists and their style of working were helpful and supportive.
- There was too much material covered in too few sessions; the intervention was not long enough - not enough time to start making life changes.
- The first few weeks of the group were emotionally challenging (the sessions covering creative hopelessness), which at times was helpful but at other times felt dysregulating and unhelpful.
- There should have been more explanation of the relationship between DBT and ACT skills, early in the group.
- The 6-month follow-up period felt like a long time to try and maintain recently initiated life changes, without any therapeutic input.

Four of the five treatment completers said that they would definitely recommend the group to a friend with similar issues. The fifth participant said that she would not recommend the group unless changes were made in line with her feedback.

7.4 Discussion

7.4.1 *Study Findings*

A small number of previous studies have suggested that ACT might be an effective intervention for PD, but prior to this study it had not been tested with a heterogeneous, post-DBT PD sample. This study was designed to begin the process of developing an ACT-based group intervention for people with poor personality functioning across PD diagnoses, who had graduated from DBT but were still reporting difficulties in terms of symptomology and engagement in a valued life. The feasibility, acceptability and impact of the intervention were assessed post therapy, and at 6-month follow-up. Given that this is the first trial of this newly developed protocol with a highly complex and potentially risky patient group, the impact of the

intervention on the group as a whole and on individual participants will be discussed in some detail.

7.4.1.1 *Impact on Valued Action and Engagement in Life*

Given that ACT is primarily designed to bring about greater engagement in a personally meaningful life regardless of symptomology, findings from measures such as the VLQ and the FS are particularly important. There was a non-significant group mean increase in values-related action during the intervention, with this trend continuing during the follow-up period. Similarly, there was a non-significant increase in group mean engagement in life (FS); RCI calculations showed that for two participants these FS improvements reached statistical significance. Several participants verbally reported increases in valued actions such as seeking and starting voluntary employment. The researchers were concerned that these important behavioural changes were not being captured by the measures included in the study, an issue that will be discussed in Section 7.4.2.

Although these changes are modest, it appears possible that the protocol had some impact on these primary outcome measures. Given the very small sample size, however, more research is needed before firm conclusions can be drawn. There are no published PD treatment trials that have included these measures, so it is not possible to compare findings.

7.4.1.2 *Impact on Parasuicidal Behaviour and Psychiatric Symptomology*

Parasuicidal Behaviour. The original aims of the intervention were to help participants engage in a valued life while maintaining behavioural stability in terms of parasuicidal behaviour. Based on the small amount of relevant previous research, improvements in mental health symptomology were also predicted. In terms of behavioural stability, no participants made suicide attempts during the group or the follow-up period, and no one was hospitalised for psychiatric reasons during these two time periods, but two of the five completers engaged in episodic, relatively minor self-harm.

Due to the small sample size and lack of a randomised controlled study design, it is difficult to interpret these findings. Certainly, some self-harm is reported in the follow-up periods of DBT for BPD trials (e.g. van den Bosch et al., 2005), and it may be that for people with very extensive self-harm histories such as the current

participants, this is to be expected. Alternatively, it could be argued that all participants reported a 6-month self-harm free period prior to the intervention (although there is doubt about this claim for one participant), and therefore these incidents of self-harm during the study are a negative reflection on the intervention. It may be relevant that both participants who self-harmed during the study reported the minimum number of months self-harm free prior to the intervention (6 months), with participants who reported longer periods without self-harm prior to the study, fairing better. Clearly, further investigation of this important issue is required.

Psychiatric Symptomology Although ACT was not developed with the specific aim of reducing psychiatric symptomology, such reductions do often occur following ACT interventions, including those for PD. In the current study, impact on both Axis I and Axis II symptomology was inconsistent. Half of the original six participants appear to have significantly benefitted from the protocol, as evidenced by reductions in Axis I and/or Axis II symptomology. For two of these participants, some of these improvements were evident at the follow-up point. Participant 2 was unable to provide psychometric data at follow-up, but one of the therapists met with her several months later, when she reported feeling positive and more active in her life than she had done prior to participating in the group.

Two participants who completed the intervention reported deteriorations in psychiatric symptoms during or following the group. Participant 1 showed little change in psychiatric symptomology immediately post-intervention, but deteriorated during the follow-up period. At the follow-up group session she described significant, aversive life events that had occurred during the follow-up period, and her view was that she would have deteriorated more severely if she had not experienced the ACT group. There is no way of determining if this was the case, and it is possible that her deterioration was linked to the intervention rather than post-group life experiences.

It is even more difficult to be certain about the experience of participant 4, who showed significant deterioration across measures of Axis I psychopathology, and a measure of personality functioning (the SIPP), but no change in her number of PD diagnoses at follow-up. It seems likely that she under-reported her level of depression and self-harming behaviour prior to the group starting. The therapists' experience of participant 4 was that she presented in challenging ways throughout the group sessions, including episodes of dissociation, and hostility towards the therapists. It is possible, of course, that these observations represent personal reactions of the

therapists rather than being accurate representations of the patient. Although she verbally described some increases in valued action, both therapists felt that on balance that participant 4 had not benefitted from the intervention.

Possible Moderators It is important to try to identify possible moderating factors that might help distinguish between participants who benefitted and those who did not, so that those who are unlikely to benefit would not be offered the intervention in the future. With the small sample size it is not possible to test potential moderators statistically, but some possible hypotheses do suggest themselves. As was outlined above, the two participants who deteriorated both had a shorter period of time self-harm free prior to the ACT group than those who benefitted. Additionally, one of them had just a few weeks following DBT before she began the ACT group, whereas the other had several years gap between DBT and ACT, and was the only participant who had self-harmed since completing DBT, suggesting that she was no longer using her DBT skills effectively. All the patients who reported benefits had had 6 to 12 months since finishing weekly DBT sessions, a period of time that perhaps can be interpreted as sufficient to allow a habit of effective skills use to be established independent of DBT input, but not long enough for skills use to have become neglected.

Although both participants who deteriorated had an above average number of PD diagnoses (five each), this on its own did not appear to make it more likely that an individual would not benefit from the intervention, because participant 5, who had six PD diagnoses, appears to have benefitted. It could be the case that patients with different PD diagnoses require different psychotherapeutic interventions, and that the protocol tested in the study was appropriate for some diagnoses but not others. This could fit with the fact that different forms of DBT are being developed for different groups of PD presentations (Lynch & Cheavens, 2008). However, looking at specific PD diagnoses, the only one that was present for either participant 1 or 4 but not the other participants is schizotypal PD (participant 4). All the other PD diagnoses for both of these participants were also associated with participants who benefitted from the group. It is possible that the intervention was unhelpful specifically with regards to schizotypal PD, although drawing this conclusion based on the experience of just one participant would be premature.

7.4.1.3 *Feasibility and Participant Satisfaction*

In terms of feasibility, the intervention performed well. Five of the original six participants completed the group (83%), an attrition rate comparable to other pilot ACT interventions for PD (e.g. Clarke et al., in prep, reported 13% attrition from their ACT condition). Session attendance was good, with a mean attendance of 78% of the group sessions, rising to 87% for participants who completed therapy.

Although patient satisfaction was not formally assessed, in post-intervention interviews four of the five completers said that they would recommend the group to a friend with similar issues, with the fifth participant indicating that she would recommend it if some specific changes were made, such as extending the length of the group. All completing participants reported having found some aspects of the intervention helpful. This feedback, although informal, suggests that the majority of the participants were satisfied with the intervention.

The participant who dropped out indicated that she had found the creative hopelessness phase of the intervention emotionally challenging. The therapists encouraged group members to reflect on the possible short-term benefits of avoidance as a strategy, particularly in high-risk, emotionally dysregulating situations, versus the potential costs of over-reliance on avoidance, in terms of lack of engagement in a satisfying life. However, more than one participant took this to mean that avoidant strategies were no longer an option open to them, and that they 'should' be able to accept any and all personal experiences. This extreme and inaccurate construing had not been observed in previous ACT groups run by the therapists, but appeared to be common in this patient group.

7.4.1.4 *Other Study Findings*

There were few significant changes in any of the process measures included in the study, despite there being theoretical and/or prior empirical justification for their inclusion. Of the two ACT-relevant process measures, the AAQII showed no significant changes at all, with just one participant reporting a significant improvement on the CFQ. Half the sample reported a change on the SCS either pre to post-group or pre-group to follow-up.

There are various possible explanations for the relative lack of change in process measure scores. One explanation is that despite prior empirical findings to the contrary, these processes (e.g. CF) are not implicated in the therapeutic change

resulting from this kind of intervention. Another possibility is that some or all of these processes do play important roles, but the intervention was not effective enough to elicit measurable change. A third possibility is that the measures are either unsuitable for this patient group, or are not sensitive enough to detect any changes that did occur. Certainly, the RCI methodology is very sensitive to a measure's test-retest reliability, where the difference between excellent and very good reliability can make the difference between significant and non-significant results.

The majority of outcome trials for psychosocial treatments for PD have focused on one PD diagnosis (usually BPD), with other co-morbid PDs rarely being reported. This perpetuates the impression that the norm is for PDs to occur independently of each other. Contrary to this impression, the participants in this study all had lifetime prevalence diagnoses of several PDs, with the majority having diagnoses from more than one PD cluster. These findings provide support for the view that there is likely to be core personality pathology that is not diagnostic category-specific, and fits with the move in DSM-V to a more dimensional approach to PD diagnosis.

In keeping with the findings from published DBT intervention trials for BPD (e.g. McMain et al., 2009), the participants in this study all reported significant, on-going psychiatric symptomology at baseline, despite having had a minimum of 12 months full-programme DBT, and having engaged sufficiently well with DBT that they had managed to successfully address their self-harming behaviours. This is not to suggest that these prior experiences of DBT had had no impact on psychopathology, but rather that DBT targeting parasuicidal behaviours had not been as effective at eliminating psychiatric symptomology. This issue will be discussed more fully in Chapter IX, Section 9.1.2. It is also possible that without DBT input, patients were deteriorating, a possibility that could be explored by collecting data over several months prior to an ACT intervention.

At baseline, participants scored highly on measures of two ACT-relevant processes; psychological flexibility and CF. These findings are consistent with previous study findings that link these processes with poor personality functioning, including those from Chapter V of this thesis. It is interesting that such high levels of inflexibility and CF were found in a sample of DBT graduates, an intervention that emphasises mindfulness practice and other acceptance-focused skills, and might have been expected to positively impact these processes. However, it was observed by the

study therapists that the group participants consistently favoured avoidance-focused DBT skills, and that the majority appeared to be using mindfulness practice as an avoidance strategy, by focusing their attention on a picture, for example, in order to try to 'block out' upsetting thoughts or feelings.

7.4.2 *Methodological Limitations*

This study was designed as an initial treatment development trial, and as such was limited in several ways. It was deliberately small in scale, an appropriate design decision when trialling a new treatment with a potentially high-risk patient group. However, the original plan was to run two consecutive groups to recruit at least 10 participants. Following the first group it was decided to modify the treatment protocol to the extent that the data from further groups could not be combined with the current data. This resulted in the current study having a smaller than planned sample size, rendering it difficult to detect group effects, and the need for caution when generalising from the findings.

As there was no control group or randomisation of participants, it is possible that changes observed in the participants could be the result of factors independent of the intervention, such as life events, or factors not specific to the intervention, such as therapeutic alliance. The study findings would therefore require replication in a larger, controlled trial for firm conclusions to be drawn.

A further limitation was the reliance on self-report measures to assess change (with the exception of the SCID-II). Although this is no different to the majority of psychotherapy outcome trials, self-report measures are vulnerable to demand bias and inaccuracy of responding, as well as participant fatigue or boredom when completing a large battery of tests. In this study, steps were taken to address these possible concerns. Participants were identified by an ID number on questionnaires rather than name, and completed the questionnaires at home, in an attempt to minimise potential demand bias. To reduce the possibility of fatigue or boredom, they were also told that they could take up to 5 days to complete each set of psychometrics.

The study measures were also limited in the sense that group members verbally reported making positive behavioural changes such as starting voluntary work, which were not necessarily reflected in scores on the study questionnaires. The VLQ, which is designed to measure change in valued activity has content that is

somewhat general and abstract in nature, and does not allow the recording of information about specific activities. Although positive shifts in VLQ action scores were seen for the group, it is possible that a measure measuring the frequency of specific behaviours might more accurately reflect the behavioural changes seen in the group.

A further methodological limitation in relation to measures is that the two VLQ variables (valuing/current importance and valued action) were both treated as outcome measures in this study. It might be more in keeping with ACT theory to use the valuing variable as a process measure, whilst retaining action in relation to values as an outcome measure.

Finally, in relation to measurement issues, it is possible that the lack of changes in score on the process measures could be a result of their not being suitable for use with this specific patient group. However, this seems unlikely, particularly in the case of the CFQ, which was shown in Chapter V to correlate highly with personality functioning. Similarly, although the AAQII has not been used in published studies involving PD samples, the earlier version of the same measure, with which the AAQII correlates very highly, has repeatedly been seen to perform well in studies involving PD samples.

7.4.3 *Protocol Limitations*

There were several possible limitations to the intervention protocol. Based on feedback from participants and the experience of the therapists, it seems likely that the protocol had too few sessions, and covered too much material in those sessions. It also seems likely that addressing creative hopelessness, and specifically the issue of acceptance versus avoidance, in the way it was addressed, could be improved. It is possible, with changes made to some aspects of the intervention that more consistently positive outcomes might be attained.

Of course it is also possible that even with such modifications, this kind of protocol will prove to have only limited benefits for this patient group. This might be due to a fundamental design aspect of the protocol, for example, that it is based on a group intervention. This particular example seems unlikely, however, as the majority of outcome trials for people with PD diagnoses are based on group interventions.

Alternatively, it might be the case that this intervention, designed as it is to address general processes such as CF, has a positive impact on some shared aspects of poor personality functioning, but that different PD diagnoses require specifically tailored interventions to have a greater impact. However, if this were the case, the fact that all of the participants in the study had several different PD diagnoses would suggest that they would each require a number of different psychosocial interventions, in addition to the 12 months or more DBT they have already received. It might therefore be reasonable to test an improved, general intervention in the first instance.

7.4.3 *Implications and Future Research*

The results from this initial treatment development trial suggest that an ACT-based group intervention for a post-DBT, heterogeneous PD sample is feasible, acceptable to patients, and may have some positive impact on engagement in valued life activities and symptomology. However, the findings were mixed, with some participant apparently not benefitting. It would be usual practice to follow a small scale, open trial of this nature, with a larger, controlled trial, to test the protocol under more illuminating conditions (Rounsaville et al., 2001). However, it seems likely that the protocol could be improved. For this reason, a second small, open trial evaluating a modified version of the protocol is the most logical next step. Changes might include:

1. Adding more sessions to the protocol.
2. Removing or simplifying some content.
3. Changing the way creative hopelessness, acceptance and avoidance are addressed.
4. Adding some top-up sessions during the follow-up period.
5. Bridging the gap between DBT and ACT more effectively.
6. Using a checklist of life activities to try to capture the behavioural changes identified in group sessions.
7. Gathering independent information about patient self-harming behaviour prior to the intervention commencing.
8. Adding a baseline period prior to the start of the group, to examine patient stability, post-DBT.

7.4.4 *Summary*

This initial treatment development trial of an ACT-based group intervention for post-DBT PD patients elicited mixed results, with several modifications to the protocol suggesting themselves as possible ways to improve the impact of the intervention. Therefore, Chapter VIII outlines a second, similar trial, examining the impact of the modified protocol.

CHAPTER VIII

Study 5: A Second Uncontrolled Pilot Development Trial of an ACT-based Group Intervention for Post-DBT Patients with Poor Personality Functioning

8.1 Introduction

Chapter VII outlined the rationale for a post-DBT ACT intervention for patients with PD diagnoses with continuing psychological difficulties and limited engagement in a valued life. The results of an uncontrolled pilot trial of this intervention were mixed, with little change for some participants on study measures. This chapter therefore reports a second uncontrolled trial of a modified version of the protocol, for post-DBT patients with poor personality functioning across PD diagnostic categories. It was hypothesised that there would be positive changes, pre to post-treatment, on study measures of engagement in life, activity, and Axis I and II symptomology.

8.2 Method

8.2.1 *Design*

The study was based on an uncontrolled, pre-post design. The IV was a 24-week ACT-based group intervention (plus psychiatric treatment as usual); the DVs being the same measures as those in the first trial. Two groups were run, sequentially. An additional data collection point was added 3-months prior to the intervention starting, as a means of determining how stable participants' scores on study variables were, post-DBT. This was seen as important, given that participants' scores on measures of psychopathology in the previous study indicated high levels of psychiatric difficulties, pre-ACT, despite all participants being DBT graduates. This raised the possibility of participant deterioration since finishing DBT.

8.2.2 *Participants*

All ethics approval, inclusion and exclusion criteria, and recruitment arrangements were identical to those in the previous study. Participants were recruited between August 2010 and July 2011. Given concerns about the accuracy of

self-report information about recent parasuicidal behaviour in the previous study, in the current study, these data were verified using the NHS electronic clinical notes system.

Figure 8.1 shows the flow of participants through the trial stages. Of the 15 people referred into the study, 10 commenced therapy (five in each of two groups). One dropped out of the second group after the first session.

Figure 8.1 *Flow Chart of Participant Recruitment, Assessment, and Treatment*

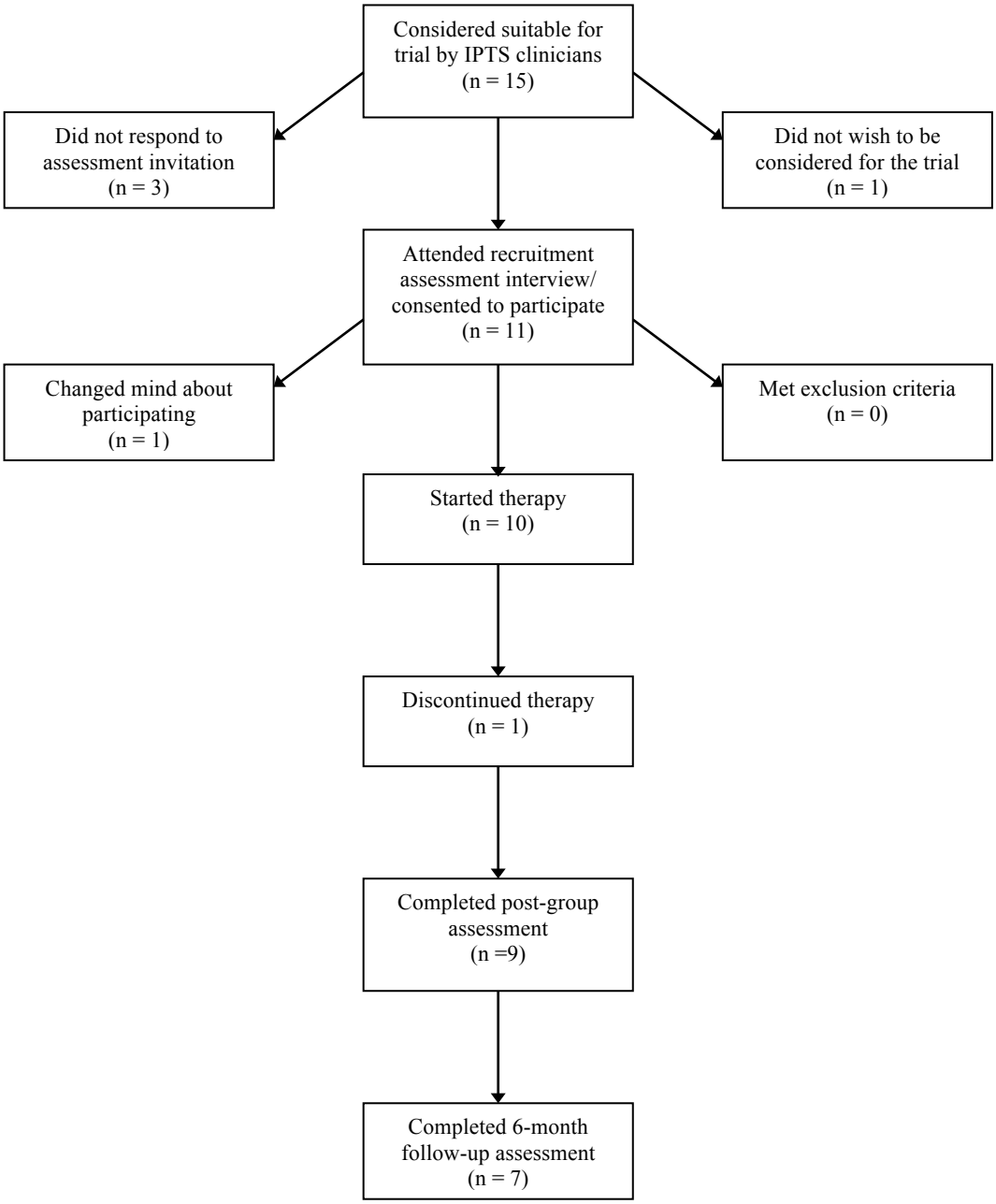


Table 8.1 contains demographic information and baseline data regarding self-harm, DBT history, and current mental health status. All participants were female, and had completed DBT prior to attending the ACT group, with the duration of DBT ranging from 16 to 26 months. The interval following the end of weekly DBT sessions and joining the ACT group varied considerably (1 to 108 months). All participants had histories of self-harm, averaging 22.50 years (ranging from 3 to 43). The period free of self-harm prior to the intervention ranged from 6 to 18 months. All participants scored in the clinical range for both depression and global symptom severity.

Table 8.1

Demographics and Baseline Statistics

Demographic/Baseline statistic	N = 10
Mean age (range)	43.10 (29 – 56)
Gender (% female)	100%
% taking psychotropic medication	100%
Mean no. months of DBT (range)[1]	19.30 (16 – 26)
Mean no. months since end DBT (range)[2]	21.70 (1 – 108)
Mean number years self-harm (range)[3]	22.50 (3 – 43)
Mean no. months since self-harm[4]	12.80 (6 – 18)
% in clinical range for depression[5]	100
% in clinical range on SCL-90R GSI[6]	100

Note. 1: Number of months of DBT intervention. 2: Number of months since the end of regular DBT sessions, by the start of ACT group. 3: Number of years history of any form of deliberate self-harm, from first reported instance of self-harm, to most recent. 4: Number of consecutive self-harm free months prior to the start of the ACT group. 5: % scoring at least 14 on the Beck Depression Inventory-II. 6: % scoring at least .70 on the Symptom Check List – 90 Revised, Global Severity Index.

Table 8.2 shows baseline PD diagnosis information for participants, indicating a very high level of personality psychopathology. All except participant 4 had PD diagnoses from two different PD clusters, with four participants having PD diagnoses from all three clusters.

Table 8.2

PD Baseline Statistics

Participants (N = 10)	N = 10
1	BPD, Obsessive-compulsive
2	BPD, Dependant, Avoidant, Depressive
3	BPD, Depressive, Obsessive-compulsive, Paranoid, Passive-aggressive
4	BPD
5	BPD, Dependant, Avoidant, Schizotypal, Anti-Social
6	BPD, Dependant, Depressive
7	Dependant, Avoidant, Depressive, Paranoid
8	BPD, Paranoid, Dependant, Avoidant, Depressive,
9	Obsessive-compulsive, Passive-aggressive Avoidant, Depressive, Obsessive-compulsive, Schizoid
10	BPD, Paranoid, Dependant, Avoidant, Depressive, Obsessive-compulsive
Mean PD diagnoses (range)[1]	4.1 (1 - 7)
Group PD diagnoses: most common to least common (number of participants with diagnosis)	BPD (8), Depressive (7), Avoidant (6), Dependant (6), Obsessive-compulsive (5), Paranoid (4), Passive-aggressive (2), Schizoid (1), Schizotypal (1), Antisocial (1)

Note. [1]Lifetime prevalence of PD diagnoses based on the SCID-II.

8.2.3 *Materials*

8.2.3.1 *Measures*

All measures used were the same as those in the first treatment development trial (see Chapter VII, Section 7.2.3), with the exception of the Life Activity Schedule

(LAS) and the Client Satisfaction Questionnaire-8 (CSQ-8; Larsen, Attkisson, Hargreaves, & Nguyen, 1979). Because several participants in the previous study mentioned positive behavioural changes in group sessions, the LAS was included in the current study in an attempt to measure such changes. The CSQ-8 was included to measure participant satisfaction (see below for details of these measures).

All except the SCID-II, LAS and the CSQ were administered at four time points; 3 months prior to the start of the intervention, pre-intervention, post-intervention, and at 6-month follow-up. The SCID-II was administered pre-intervention and at 6-month follow-up, as per the previous study. The LAS was completed as part of the routine group administration procedure by participants prior to the start of each group session. The CSQ-8 was completed only once, shortly after the intervention ended.

Life Activities Schedule (LAS) See Appendix O. This idiographic measure of life activity frequency and type was developed specifically for this trial. It is an anglicised version of the Life Activities Checklist included in Lejuez, Hopko, and Hopko, (2001), a behavioural activation treatment for depression trial. The original checklist did not include a scale, so a 4-point scale was added, ranging from *Once in the last week* to *More than once a day in the last week*. The schedule covers 131 positive or enjoyable activities, (examples include "Visiting friends or having friends visit", "Listening to music", and "Doing charity work"), and has space for participants to add a further six activities. The measure yielded two variables; mean weekly activity level across all activities, and mean number of types of activity per week. No psychometric data are available for this measure.

Client Satisfaction Questionnaire-8 (CSQ-8) This is an 8-item self-report measure of client satisfaction with health services. Items include "Have the services you received helped you to deal more effectively with your problems?" and "How would you rate the quality of the service you received?" Responses are on a 4-point scale, the labelling of which varies between items. The measure yields a total score ranging from 8 to 32, with higher score indicating greater satisfaction. It has very good to excellent internal consistency (Cronbach's α ranging from .86 to .93).

ACT is primarily designed to increase valued action and engagement in life, rather than to reduce symptomology. For this reason, the VLQ, FS, and LAS were used as primary outcome measures, along with two measures of personality functioning (SCID-II and SIPP).

8.2.4 *Procedure*

Recruitment and assessment followed the same procedures as the first trial (see Chapter VII, Section 7.2.4).

8.2.4.1 *Intervention*

The following changes were made to the original ACT protocol (which was outlined in Section 7.2.4.4).

Number of sessions The number of weekly sessions was increased from 20 to 24 because participants from the first trial reported they would have benefitted from more time to make valued life changes. In addition, 1-month and a 3-month follow-up sessions were added, again following participant feedback.

Content of sessions Some changes were made to the content of the ACT group protocol. More detailed coverage of the links between DBT and ACT was included at the beginning of the group, to aid participants' transition from one approach to the other. A small amount of content, both didactic and experiential, was removed from several sessions, to allow more time to address the remaining material. No topics were completely removed from the protocol. Additionally, the content addressing RFT and experiences of self was simplified.

Some participants from the previous trial reported finding the creative hopelessness aspect of the intervention—a feature of the early stages of many ACT mental health protocols—anxiety provoking. This was because their primary coping strategy (avoidance) was being singled out as potentially unhelpful, prior to any additional skills or coping strategies having been taught. Although the "control and avoidance is the problem" message inherent in many ACT protocols had already been modified to be less disturbing to this patient group (see Chapter VII, Section 7.2.4.6 for details), further changes were made. Specifically, the order of the early sessions was altered so that a new ACT skill (as it was referred to in the group)—cognitive defusion—was introduced and practiced, prior to any exploration of the potential costs of avoidance. In this way, the willingness and ability to allow private experiences such as thoughts and emotions to be present, was introduced as something that all participants were already familiar with (through repeated defusion practices).

Finally, as indicated above, because the group members were familiar with the idea of learning and utilising new skills (from DBT), the main capacities associated with the six ACT processes were conceptualised in this way in the modified protocol. They were added to the DBT skills on the diary card completed by participants on a daily basis, and as each new ACT skill was taught, group members were encouraged to practice and use them alongside their DBT skills. The ACT skills were described on the diary card as follows:

- Cognitive defusion
- Clarifying values
- Practicing willingness
- Taking valued action
- Connecting with YOU (not your mind) who can notice and choose
- Noticing sweet moments

8.2.5 *Analysis Plan*

The same analysis plan as the first trial was used, based on non-parametric group statistics and individual RCI calculations, but with a small number of modifications. Firstly, possible group differences between 3-month pre- and immediately pre-group scores on all measures were investigated in Stage Two, to assess the stability of scores, post-DBT and prior to ACT. Secondly, effect sizes were calculated. Finally, two additional measures (the LAS and the CSQ-8) were included.

8.3 Results

8.3.1 *Stage 1. Preliminary Analysis and Participant Characteristics*

There was less than 1% data missing from the psychometric measures, so missing items were replaced by the sample mean (Tabachnick & Fidell, 2001). One participant dropped out after the first session. She completed 3-month pre- and

Table 8.3

Means and Standard Deviations for All Study Measures

	Mean (SD)			
Measures	3-month pre N = 10	Pre-group N = 10	Post-group N = 9	Follow-up N = 7*
FS	33.44 (11.30)	31.11 (10.86)	35.44 (8.50)	35.14 (9.87)
Activity	-	.39 (.31)	.39 (.32)	.39 (.22)
Activity types	-	22.11 (10.35)	27.33 (17.29)	26.78 (11.13)
VLQ CI	6.05 (2.63)	5.33 (1.29)	6.71 (2.32)	7.90 (1.63)
VLQ CA	4.94 (2.40)	4.47 (1.65)	4.62 (1.80)	7.36 (1.08)
SCID-II	-	2.33	-	2.22
SIPP S-con	4.45 (.72)	4.27 (.82)	4.55 (.83)	4.42 (.99)
SIPP Social	5.48 (1.05)	5.19 (.95)	5.49 (1.02)	5.42 (1.38)
SIPP Ident	3.17 (.76)	3.08 (.63)	3.14 (.59)	3.12 (.66)
SIPP Relat	3.62 (.98)	3.43 (.79)	3.61 (.91)	3.48 (.74)
SIPP Respon	4.25 (.85)	4.13 (.84)	4.21 (.96)	4.14 (.79)
DSHI	0 (0)	0 (0)	1.22 (2.22)	1.14 (1.95)
BDI-II	33.11 (13.68)	35.78 (11.37)	33.33 (8.17)	27.57 (12.39)
SCL-90-R GSI	1.75 (.73)	1.77 (.77)	1.94 (.73)	1.68 (.91)
CFQ	59.67 (8.94)	60.89 (11.60)	62.56 (11.17)	56.14 (11.45)
AAQII	33.00 (4.50)	32.00 (5.27)	34.56 (6.35)	32.71 (3.82)
SCS	14.54 (2.84)	14.22 (2.95)	14.69 (3.66)	14.86 (4.39)
CSQ	-	-	28.67 (2.74)	-

Note. FS = Flourishing Scale; Activity = LAS mean number of activities; Activity types = LAS mean number of types of activity; VLQ CI = Valued Living Questionnaire Current Importance; VLQ CA = Valued Living Questionnaire Current Action; SCID-II = Structured Clinical Interview for DSM-IV Axis II Disorders; SIPP S-con = self control domain; SIPP Social = social concordance domain; SIPP Ident – identity integration domain; SIPP Relat = relational functioning domain; SIPP Respon = responsibility domain; BDI-II = Beck Depression Inventory; SCL-90 GSI – Symptom Checklist-90 Global Severity Index; DSHI = Deliberate Self-Harm Inventory; CFQ = Cognitive Fusion Questionnaire; AAQII – Acceptance and Action Questionnaire 2nd Version; SCS = Self Compassion Scale.

pre-group measures only. Patients attended 18 of the 24 sessions, on average. This figure rose to a mean of 19 for the treatment completers. Table 8.3 shows group means scores on all study variables at the four data collection time points. Two sets of data were missing at follow-up, due to one participant having just given birth, and another being unavailable due to her partner being seriously ill. Pre-intervention, this sample showed poorer functioning on all SIPP personality domains compared with the PD sample norms provided by the authors of the measure (Andrea, personal communication). In terms of Axis I pathology (BDI-II and SCL-90_R GSI), this sample was comparable to the high-risk BPD samples in DBT and CBT for BPD trials (e.g. Davidson et al., 2006; McMain et al., 2009).

8.3.2 *Stage Two. Group Statistical Analyses and Effect Sizes*

8.3.2.1 *Group Statistical Analyses*

Friedman's ANOVAs were conducted at pre- post, and follow-up, on all primary outcome variables except the SCID-II. All participants completing the intervention were included in the analyses (n = 9), with the post-intervention data from participants 6 and 7 being carried forward to the follow-up data point, (following recommendations by Hollis and Campbell, 1999), because they completed the intervention sessions but were not able to provide follow-up psychometric data.

Table 8.4 shows there was a significant increase in valued action across the three data points included in the analysis, which remained significant following Bonferroni correction for multiple comparisons (p = .05/10 = .005). None of the other changes were significant.

A Wilcoxon signed-rank test yielded no significant difference between the two data collection points for the SCID-II, indicating no significant difference between numbers of PD diagnoses from pre-intervention to 6-month follow-up, (z = -.28, p = .78).

Based on the Friedman Test results, post-hoc Wilcoxon signed-rank tests were carried out on the VLQ current action data. Pre- to post-intervention, there was no significant change (z = 0, p = 1.0). Post-intervention to 6-month follow-up there was a significant increase in valued action (z = -2.37, p = .02).

Table 8.4

Friedman ANOVAs for Repeated Measures at Pre, Post, and Follow-Up

	n = 9		
Measures	*chi-sq*	*df*	*p*
FS	1.75	2	.42
VLQ CI	2.36	2	.31
VLQ CA	11.53	2	.003**
SIPP S-con	2.57	2	.28
SIPP Respon	2.87	2	.87
SIPP Ident	.29	2	.87
SIPP Relat	2.57	2	.28
SIPP Social	.29	2	.87
Activity	.06	2	.97
Activity type	2.97	2	.23

Note. **Significant at the .05/10 = .005 level; FS = Flourishing Scale; VLQ CI = current importance (valuing); VLQ CA = current action; SIPP S-con = self control domain; SIPP Social = social concordance domain; SIPP Ident – identity integration domain; SIPP Relat = relational functioning domain; SIPP Respon = responsibility domain; Activity = LAS mean number of activities; Activity types = LAS mean number of types of activity.

8.3.2.2 *Effect Sizes*

It is common practice to calculate effect sizes for changes in group means, where non-significant results could result from the study being statistically underpowered. (e.g. Goldstein et al., 2007.) Based on Cohen's (1988) guidelines regarding effect sizes, it can be seen in Table 8.5 that the changes in several of the study variables are of medium or large size. From pre to post-intervention, there were medium or large effect sizes in the direction of better functioning, for flourishing in life, activity type, valuing, self-control, and social concordance.

Table 8.5

Effect Sizes

Measures	3-month-Pre	Pre-Post	Pre-F-UP
	ES	ES	ES
FS	.29*	1.08	.52
Activity	-[1]	0.00	0.00
Activity type	-[2]	.67	.71
VLQ CI	.45*	.91	2.82
VLQ CA	.62*	.14	2.51
SCID-II	-	-	.22
SIPP S-con	.49*	.88	.32
SIPP Social	1.10*	.65	.42
SIPP Ident	.21*	.17	.11
SIPP Relat	.39*	.32	.08
SIPP Respon	.45*	.28	.03
DSHI	0	1.26*	1.58*
BDI-II	.40*	.36	.84
SCL-90 GSI	.05*	.31*	.22
CFQ	.12*	.50*	.51
AAQII	.42	.83*	.44*
SCS	.14*	.29	.36

Note. *Effect in the direction of poorer functioning; [1]and[2]: activity data was not collected prior to the intervention.

FS = Flourishing Scale; Activity = LAS mean number of activities; Activity types = LAS mean number of types of activity; VLQ CI = Valued Living Questionnaire Current Importance; VLQ CA = Valued Living Questionnaire Current Action; SCID-II = Structured Clinical Interview for DSM-IV Axis II Disorders; SIPP S-con = self control domain; SIPP Social = social concordance domain; SIPP Ident – identity integration domain; SIPP Relat = relational functioning domain; SIPP Respon = responsibility domain; BDI-II = Beck Depression Inventory; SCL-90 GSI – Symptom Checklist-90 Global Severity Index; DSHI = Deliberate Self-Harm Inventory; CFQ = Cognitive Fusion Questionnaire; AAQII = Acceptance and Action Questionnaire 2nd Version; SCS = Self Compassion Scale.

There were similar sized positive effects for post-follow-up changes in valuing, valued action, depression, general psychiatric symptomology, fusion and psychological flexibility. From pre-intervention to follow-up there were medium or large positive effects in flourishing, activity type, valuing, valued action, depression and defusion. Deliberate self-harm showed a large effect in the direction of more incidents of self-harm, as there were no occurrences of self-harm at the pre-group time point. Both fusion and psychological flexibility were associated with medium to large effects in the direction of poorer functioning, pre-post group, with this direction being reversed for both variables post-group to follow-up, resulting in a positive, medium-sized effect for changes in fusion, pre-follow-up, and a negative medium-sized effect for changes in flexibility, pre-follow-up.

8.3.2.3 *3-Month Pre-Intervention Data*

There were no significant differences between mean ranked scores 3 months prior to the intervention and immediately pre-intervention on any of the study measures, as tested using Wilcoxon signed-rank tests. However, for the majority of the study variables, effect size calculations indicated group mean changes in the direction of poorer functioning in the 3 months prior to the start of the ACT intervention (Table 8.5), suggesting a general deterioration in functioning. These changes ranged from negligible in size (SCL-90-R), to very large (social concordance). The only exceptions were the AAQII, which yielded a medium-sized effect in the direction of better functioning, and the DSHI (the latter due to self-harm in the months prior to the intervention being an exclusion criterion).

8.3.3 *Stage Three. Individual Participant RCIs*

Individual participant RCIs were calculated for the following measures: SIPP, FS, BDI-II, SCL-90 GSI, CFQ, AAQ-II, and the SCS, from pre- to post-intervention, and from pre-intervention to follow-up (see section 7.3.3 for details of this approach). RCIs were not calculated for the LAS, SCID-II, DSHI, or VLQ, as the psychometric data needed to carry out these calculations were not available for the form of these measures used in the study. The results are displayed in Table 8.6, with those of the primary outcome measures also displayed graphically in Figure 8.2.

The majority of participants showed no statistically reliable changes between pre-group and post-group or follow-up, on the majority of measures. However, it can be seen in Figure 8.2 that for the FS, the majority of participants are above the line of no change on the pre to post graph, indicating changes in score in the direction of better functioning, which were not large enough to be considered reliable. A similar pattern can be seen on the pre-post group graphs for the SIPP self-control, identity integration, and social concordance subscales. This pattern fits with the medium to large effect sizes reported for these variables other than identity integration, in Section 8.3.2.2. The other SIPP graphs indicate a less consistent pattern, with participants variously moving in the direction of better functioning, showing little or no change, and in a minority of cases, moving in the direction of poorer functioning.

Participant 10 showed consistent, reliable and clinically significant change across many of the study measures at both the post and follow-up time points. Participants 6 and 8 showing reliable and clinically significant changes on more than one outcome measure, post-group. Participant 5 showed deterioration across more than one measure, both post-group and at follow-up.

Table 8.6

Reliable and Clinically Significant Individual Changes

	Post (n = 9)				Follow-up (n = 7)			
Measures	Recovered	Improved	Same	Deteriorated	Recovered	Improved	Same	Deteriorated
FS	0	2	7	0	0	0	7	0
SIPP S-con	0	2	7	2	1	0	5	1
SIPP Social	1	0	8	0	1	0	6	0
SIPP Ident	0	2	6	1	0	1	6	0
SIPP Relat	0	1	8	0	0	1	6	0
SIPP Respon	1	0	8	0	0	0	7	0
BDI-II	0	3	5	1	2	0	4	1
SCL-90 GSI	0	0	8	1	0	1	5	1
CFQ	0	0	9	0	0	1	6	0
AAQII	0	0	8	1	0	0	6	1
SCS	1	0	8	0	0	0	7	0

Note. FS = Flourishing Scale; SIPP S-con = self control domain; SIPP Social = social concordance domain; SIPP Ident – identity integration domain; SIPP Relat = relational functioning domain; SIPP Respon = responsibility domain; BDI-II = Beck Depression Inventory; SCL-90 GSI – Symptom Checklist-90 Global Severity Index; CFQ = Cognitive Fusion Questionnaire; AAQII – Acceptance and Action Questionnaire 2nd Version; SCS = Self Compassion Scale.

Figure 8.2

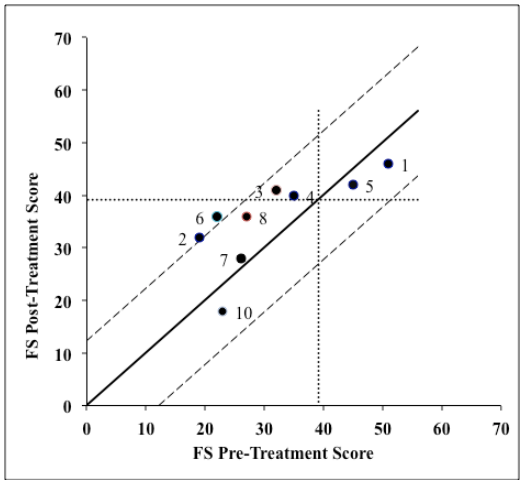
Graphs of Reliable and Clinically Significant Change for Each Participant

Legend

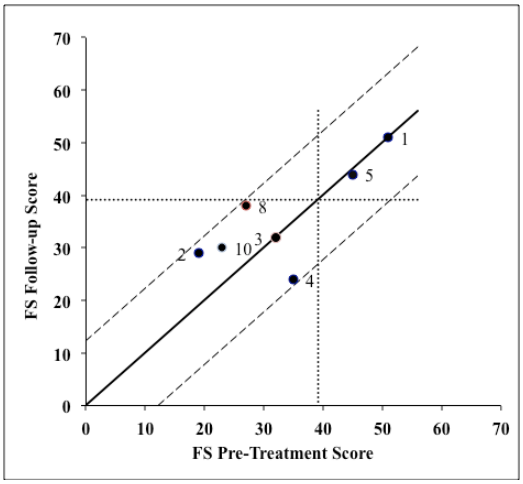
— Line of no change ----- Limits of reliable change ….. Clinical cut-off

Flourishing Scale

Pre-Post

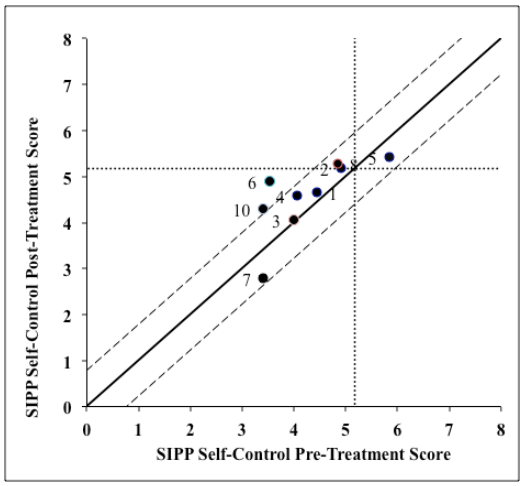


Pre-Follow-Up

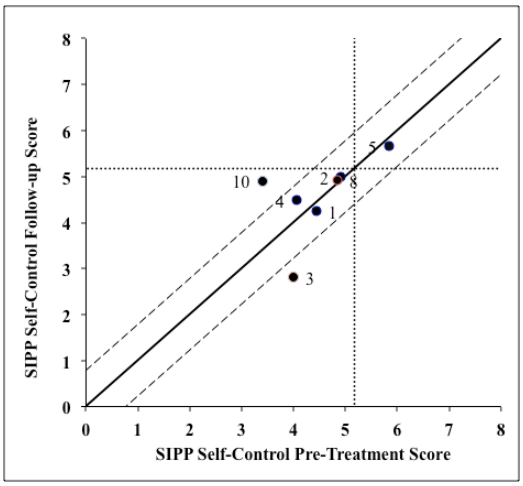


SIPP Self Control Subscale

Pre-Post

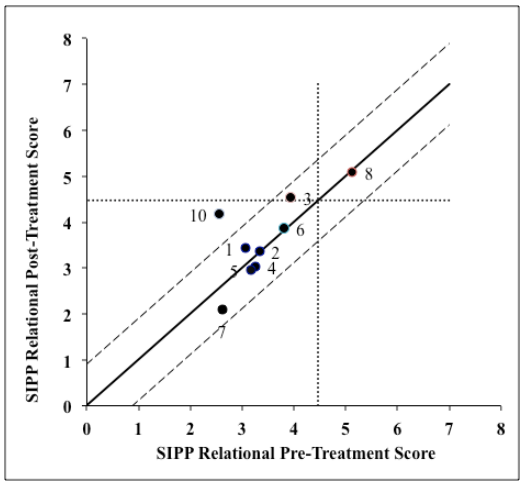


Pre-Follow-Up

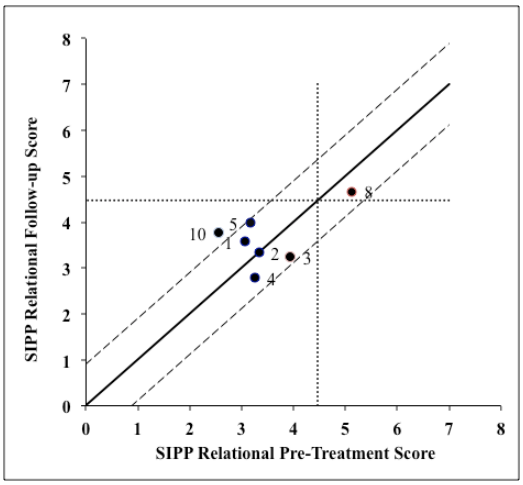


SIPP Relational Functioning Subscale

Pre-Post

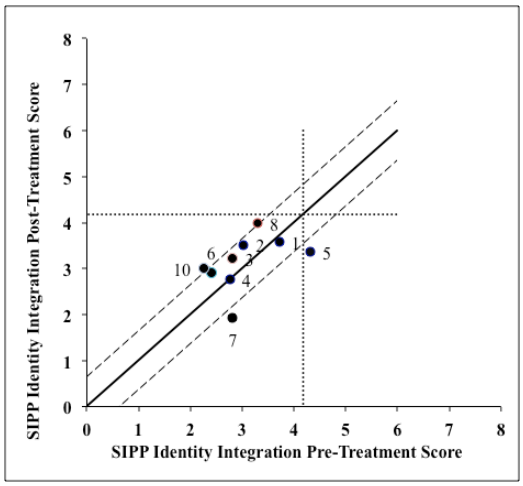


Pre-Follow-Up

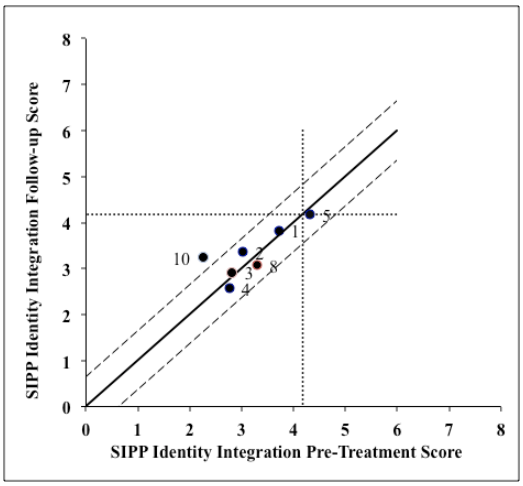


SIPP Identity Integration Subscale

Pre-Post

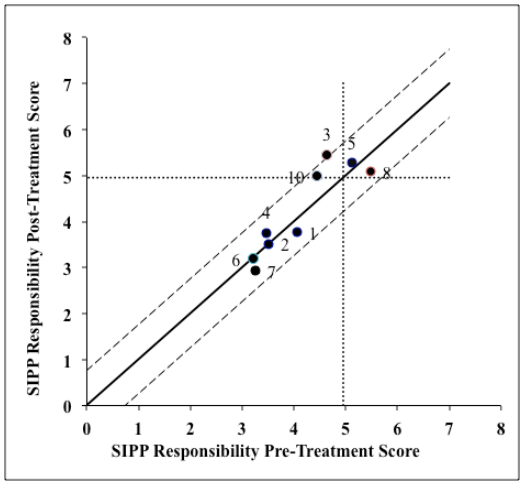


Pre-Follow-Up

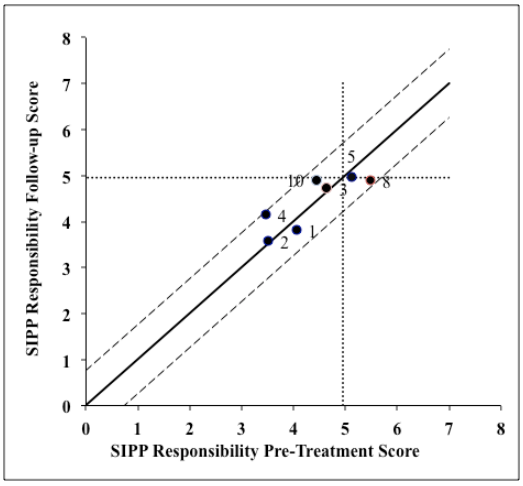


SIPP Responsibility Subscale

Pre-Post

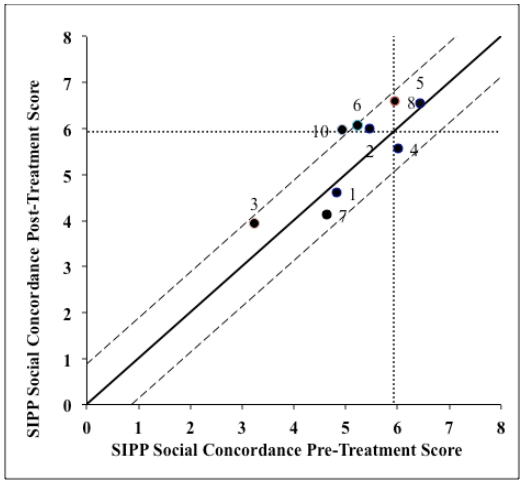


Pre-Follow-Up

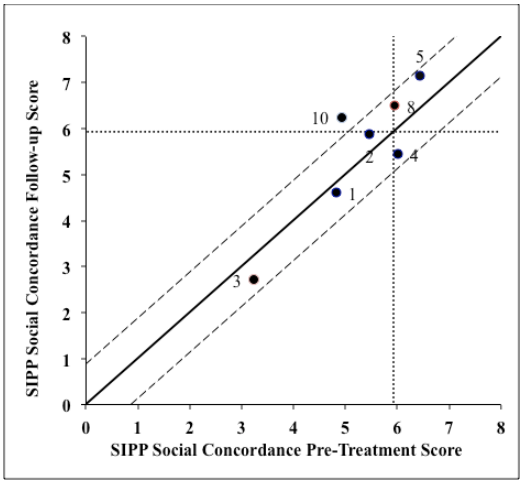


SIPP Social Concordance Subscale

Pre-Post



Pre-Follow-Up



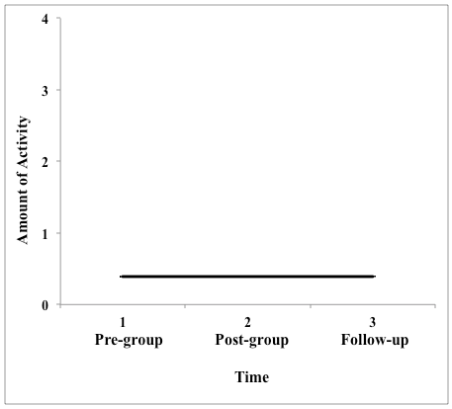
8.3.4 *Stage Four. Data Not Analysed Using the RCI Methodology*

The LAS, VLQ, SCID-II and DSHI data could not be analysed using the RCI methodology, and are presented in tabular and graph form below. Figure 8.3 shows that amount of activity, as measured by the LAS, did not change pre to post-group, or pre-group to follow-up. However, the different types of activities group members engaged in did increase from pre to post-group, with much of that increase being maintained at 6-month follow-up, suggesting a broadening of participants' behavioural repertoires. The amount of activity specifically associated with personal values (measured by the VLQ), did increase for participants – marginally from pre to post-group and substantially from post-group to follow-up. This latter increase in valued action followed an increase in valuing (current importance of valued life domains) during the intervention; an increase that continued during the follow-up period.

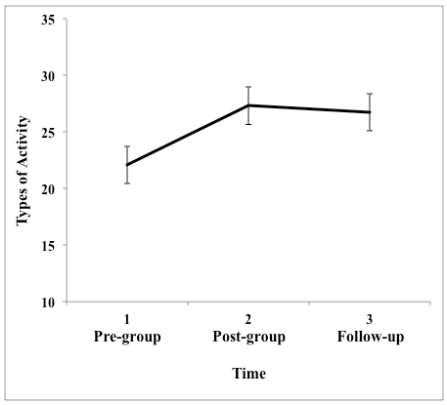
Figure 8.3

Group Means for LAS, VLQ, SCID-II, and DSHI, with Standard Error Bars

*LAS amount of activity**

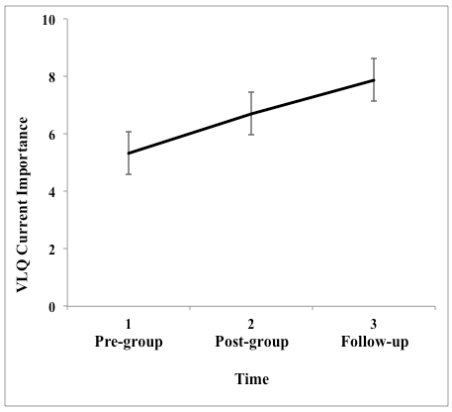


LAS types of activity

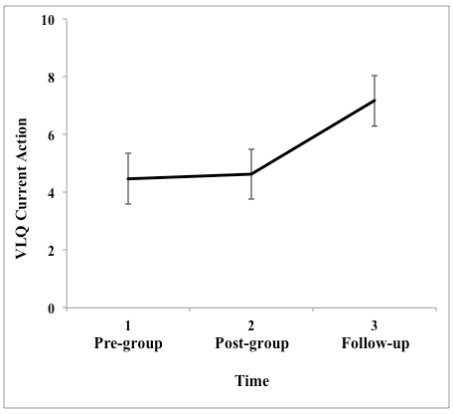


*Each activity was rated on a 0 – 4 scale as follows:
0 = Not at all in the last week; 1 = *Once in the last week; 2 = More than once in the last week; 3 =Every day in the last week; 4 = More than once a day in the last week*. The amount of activity variable was based on the mean of this rating across all activities on the measure.

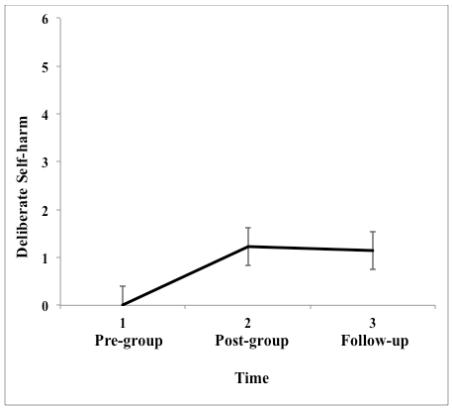
VLQ Valuing



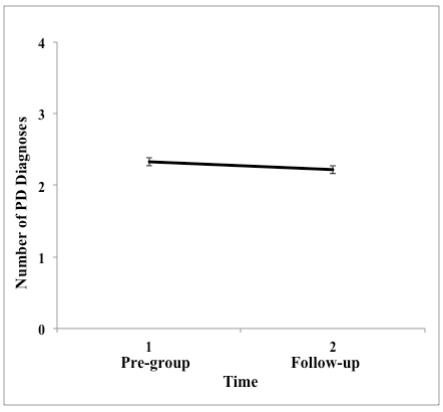
VLQ Action



DSHI



SCID-II



Changes in number of PD diagnoses as determined by the SCID-II are displayed in Table 8.7. Three of the seven participants for whom there were follow-up SCID-II data available, showed improvements, two deteriorated, and two showed no change. For some participants (e.g. 9 and 5), their change in number of PDs was consistent with changes on other measures. However, a notable exception to this was Participant 4, who showed a mixed set of results on the other study variables, but significant deterioration in terms of PD diagnoses, an issue that will be addressed in the discussion section.

Table 8.7

Changes in PD Diagnoses Based on 12-Month SCID-II Prevalence

Participant	Pre-therapy	6-month follow-up
1	1	0
2	3	2
3	3	3
4	1	4
5	1	2
6	3	-
7	1	-
8	3	3
10	5	2
Mean no. PDs	2.3	2.2*

Note. *Mean based on data from participants 6 and 7 being carried forward from pre-intervention

Changes in participant self-harming behaviour are outlined in Table 8.8. No participant made a suicide attempt either during the intervention or the follow-up period. The majority of participants engaged in no self-harming behaviour either during the group or the 6-month follow-up period. Participants 4, 5, and 7 all reported a small number of self-harm incidents during the life of the group, with Participants 4 and 5 also reporting some self-harm during the follow-up period. None of the

participants required medical attention as a result of these self-harm incidents, and using the criteria drawn up for an on-going RCT involving participants with PD diagnoses (Lynch et al., 2011 – 2016), all incidents in the current trial would be categorised as either adverse events or adverse reactions, the two least serious categories of incidents.

Table 8.8

Participant Self-Harming Behaviour

Participant	During therapy	During 6-month follow-up
1	0	0
2	0	0
3	0	0
4	1 x cut	2 x cut, 1 x skin pick, 1 x head bang
5	3 x cut, 3 x head bang	1 x cut, 3 x head bang
6	0	0
7	4 x punched wall	0
8	0	0
10		0

Note. Where no follow-up psychometric data on self-harm were available, electronic health records were accessed for this period to assess whether any self-harm occurred.

8.3.4.1 *Participant satisfaction*

Participant satisfaction was assessed using the CSQ-8, administered within two weeks of the end of the intervention. The mean total score was 28.7, out of a possible 32, indicating very high levels of satisfaction with the intervention, in terms of quality of the service and the extent to which it met participants' needs. This result compared favourably with the mean score of 26.9 from the 10 DBT and CAT treatment completers between May and September 2012 at the same PD treatment clinic where the current trial was based.

8.4 Discussion

8.4.1 *Study Findings*

Chapter VII reported an initial test of a group-based ACT intervention for DBT graduates with poor personality functioning, which indicated that the intervention was both feasible and acceptable to the patient group. Some increases in valued action and engagement in life were reported, but overall, the outcomes were mixed, particularly in terms of change in psychiatric symptomology. Several modifications were made to the treatment protocol and the current study was designed to test the feasibility, acceptability and impact of the modified protocol, in a second, small-scale, uncontrolled trial. Given that the participants have highly complex and potentially risky presentations, both group and individual responses to the protocol will be discussed in detail.

8.4.1.1 *Impact on Valued Action and Engagement in Life*

ACT aims primarily to increase valued action and engagement in life, rather than targeting symptom reduction. The current protocol had a positive impact on valued action, as measured by the VLQ, and flourishing/engagement in life, as measured by the FS. Despite the small sample size, a significant group increase in valued action was found. The timing of this change was interesting in that although valuing (the other VLQ variable) increased pre- to post-intervention, valued action did not increase at the same time. Rather, what appears to have happened is that a shift towards a greater connection with personal life values preceded the significant increase in valued action reported at the 6-month follow-up point. The design of the study and small sample size does not allow for any possible causal relationship between increases in valuing and action to be tested.

No statistically significant group effects were detected other than for valued action, but because the study was underpowered due to the small sample size, effect sizes were calculated. These showed medium to large positive effects, both post-intervention and at 6-month follow-up, for changes in engagement in life (FS), and number of types of positive activities (LAS), as well as valuing (VLQ).

Although values-related activity increased significantly, there was no change in general levels of positive activity, as measured by the LAS. The protocol did target personal values-related action, specifically, but it is unexpected that changes in

valued action were not reflected at all on the LAS amount of activity variable. Based on the effect size calculations, what did increase pre- to post-intervention was the number of different kinds of positive activities participants engaged in, an increase that was maintained at follow-up. So, although participants did not increase the overall amount of positive activities they engaged in, they did broadened their repertoire of such behaviours. None of the published studies using ACT to treat PD included measures of values or activity, so it is not possible to compare these findings with those from other relevant research. There was a very large positive effect for engagement in life (FS), pre- to post-intervention and, although this decreased somewhat over the follow-up period, the pre-intervention to follow-up effect was still medium sized. The only one of these engagement in life and values-related variables for which RCIs could be calculated was the FS, with the results indicating that two participants showed reliable changes, pre to post-intervention, with the majority of participants showing non-reliable positive changes.

In summary, the ACT group intervention was associated with large increases in values-related activity, types of activity, and engagement in life, with these improvements for the most part being sustained or even increasing at 6-month follow-up. Given that ACT is primarily designed to bring about these kinds of changes, rather than targeting reductions in symptomology, these are highly promising results.

8.4.1.2 *Parasuicidal Behaviour and Psychiatric Symptomology*

Parasuicidal Behaviour This intervention was designed to increase valued living while maintaining progress made through DBT in relation to parasuicidal behaviour. No participants made suicide attempts or were hospitalised due to psychiatric issues during the intervention or follow-up period. Three of the nine completers engaged in episodic, relatively minor self-harm, with none of these episodes requiring medical attention. As discussed in Chapter VII, Section 7.4.1.2, some re-emergence of self-harm has been reported in the follow-up period in both DBT and TAU conditions in BPD trials (e.g. van den Bosch et al., 2005), but due to differences between the various studies, it is not possible to directly compare findings. To determine whether the level of self-harm reported in the present study indicates that the ACT protocol had a negative impact, or alternatively was somewhat protective in relation to parasuicidal behaviours, requires further research, based on a randomised, controlled design with a larger sample.

Participant 4, who self-harmed during both the intervention and follow-up periods, reported that the only reason she had managed to not harm herself during the 6 months pre-intervention was because she knew that she would not be able to access ACT if she acted on her urges. It may be that potential participants require motivation to stay self-harm free beyond getting access to therapy. The other participant who self-harmed both during and after the intervention, participant 5, had been free from substance misuse for 6 months prior to the intervention, but episodic binge drinking during the life of the group appeared to have increased her risk of self-harm.

As discussed in relation to Study 4 (Section 7.4.1.2), some re-appearance of self-harm during follow-up has been reported in DBT trials for PD (e.g. van den Bosch et al., 2005), and further research in the form of an RCT would be needed to understand more fully, the impact, positive or otherwise, of ACT on self-harm for this patient group.

Psychiatric Symptomology ACT was not developed primarily to reduce psychiatric symptoms, but given that previous ACT for PD protocols have positively impacted such variables, it was predicted that this protocol would have a similar effect. There were no significant changes in group means on either Axis I or Axis II psychopathology measures. Effect size calculations indicated a small positive effect for the BDI-II (depression), pre to post-intervention, and a small negative effect for the SCL-90-R (general psychiatric symptomology) over the same time period. However, both measures were associated with large positive group effects at 6-month follow-up. For the BDI-II this was large enough to yield a large, positive pre-group to follow-up effect size. The RCI analyses support these findings, with the majority of participants, particularly during the follow-up period, reporting changes on the BDI-II and SCL-90-R in the direction of better functioning. For a third of the sample, these changes were statistically significant.

There were no such increases in Axis I functioning in the follow-up period in the test of the earlier version of this protocol. The improvement with the current version might be due to changes in the content of sessions, and/or the addition of two sessions during the follow-up period, designed to maintain the use of ACT and DBT skills to support on-going engagement in life, and to address any participant-identified problems.

These changes in Axis I symptomology are smaller than those on the same measures reported by Clarke et al. (in prep), a trial of ACT for treatment resistant

mental health problems including PD. However, their sample scored substantially lower on both measures prior to therapy (mean score of 24.23 on the BDI-II compared to 35.78 for the current sample; mean score of 1.27 on the SCL-90-R GSI compared to 1.77), and only 50% of their sample had a co-morbid PD, suggesting that the current sample had significantly more severe and complex symptomology, with more modest therapeutic improvements likely to be expected. In fact, in terms of Axis I symptomology, the current sample was comparable to behaviourally unstable, BPD samples included in CBT and DBT RCTs, prior to treatment; a point that will be discussed in relation to both clinical studies, in Chapter IX, Section 9.1.2.

In terms of measures of personality functioning, effect size calculations indicated that on the self-control and social concordance SIPP domains, there were medium to large positive effects, pre to post-intervention. The remaining SIPP domains, and the SCID-II showed small positive effects. All SIPP domains showed some changes in the direction of poorer functioning during the follow-up period, but all SIPP group mean changes, pre-intervention to follow-up, remained in the direction of better personality functioning. This pattern was echoed in the RCI analyses, where there were non-significant trends in the direction of better functioning, for the majority of participants on the majority of SIPP domains. Only participant 10 showed consistent statistically reliable positive change on the SIPP domains, a result that was echoed in a reduction in her number of PD diagnoses.

Across the sample, over 40% of the participants who provided follow-up SCID-II data showed reductions in number of PD diagnoses. One participant showed an increase in PD diagnoses (participant 4). This was not consistent with her SIPP scores, where she showed no reliable negative changes. Pre-intervention, she reached the diagnostic threshold for just one PD. This did not fit with her pattern of (high) scores on other measures of psychopathology or the experience of the group therapists, who viewed her as one of the more complex participants in the trial. It is possible that she under-reported her personality symptomology in the face-to-face SCID-II assessment prior to the group.

Overall, some positive changes in both Axis I and Axis II symptoms occurred in relation to this intervention, despite the relatively severe nature of the participants' presentations, and the fact that ACT does not directly address symptomology. However, in many cases these gains were modest. Interestingly, the main improvements in Axis I symptoms occurred in the follow-up period; perhaps

indicating how difficult and slow therapeutic change can be for this patient group. It should be noted that effect size calculations indicated a pattern of deterioration in functioning in the 3 months prior to the start of the ACT intervention. Therefore, the intervention may have served to reverse this trend and to produce modest improvements on study outcome measures, a hypothesis that will be discussed further in Section 8.4.3.4.

ACT suggests itself as a good candidate for a post-DBT intervention for PD patients for several reasons, including its emphasis on engagement in a valued life, and the possibility of ACT addressing emotional avoidance (which was Linehan's original intention for a Stage II DBT intervention). It is also a transdiagostic approach, designed to impact universal psychological processes, such as CF, and as such might be particularly suitable for people with complex patterns of personality pathology. It is, however, possible that, although ACT may positively address processes that are implicated in both Axis I and Axis II psychopathology, patients might also require therapeutic input that is tailored to their particular experiences. If that is the case, the relatively modest improvements in symptomology reported here may represent the limitations of this general approach. Of course ACT protocols are often tailored to suit a particular patient group, as Gratz and Gunderson (2006) did in their ACT/DBT group for BPD patients. Whether there would be a way of meaningfully subdividing the current sample, to enable some narrowing the focus of the ACT protocol is unclear, given the very broad nature of personality symptomology of the participants. Lynch and colleagues (see Lynch & Cheavens, 2008) have successfully taken this type of approach in their treatment of chronic depression with co-morbid emotionally over-controlled PDs.

8.4.1.3 *Feasibility and Participant Satisfaction*

The intervention performed well in terms of feasibility, with nine of the ten participants completing therapy, a result that is comparable or better than those from other ACT for PD pilot trials (e.g. Clarke et al., in prep, reported an attrition rate of 13% in their ACT condition). Group attendance was good, and satisfaction with the intervention was very high. Participants rated this intervention slightly higher than other PD treatment-completers rated DBT and CAT; the therapies that are standard care for PD in the same clinic.

8.4.1.4 *Other Study Findings*

There were no significant changes in group mean scores on any of the process measures, and few participants showed reliable individual changes on these measures. Effect size calculations showed that for the ACT-relevant measures (the CFQ and the AAQII), there was a medium and large negative effect respectively, pre to post-intervention. Both measures were associated with very large positive effects sizes when comparing post-intervention and follow-up scores. If these changes were in response to the intervention, then it appears as if, when invited to consider what a personally meaningful life might look like, and then to engage more fully in such a life, participants initially became less psychologically flexible and more fused. Over time, perhaps through continued, gradual experiential engagement with these ACT-relevant processes, it appears as if a substantial positive shift occurred.

Although this interpretation is based on limited evidence and requires future investigation, it does fits with the therapists' observations of participants experiences, in that, even with the modifications made to the ACT protocol outlined in Chapter VII, Section 7.2.4.6 and Section 8.2.4.1 of this chapter, some participants were still initially anxious about making changes in their lives, and found the focus in ACT on experiencing emotions, cognitions and so on, counter to their usual coping strategies, and somewhat challenging. However, with encouragement to make flexible choices about when to engage and when to avoid, and with repeated, small practice steps, the majority of participants appeared to increase in confidence and willingness, over time. It is possible that any intervention post-Stage I DBT, that addressed engagement in life (and therefore also engagement with private experiences such as emotions), would result in similar changes on these process measures.

Based on the limited data from this pilot study, the temporary increase in CF and inflexibility does not appear to necessarily be detrimental; Participant 10, for example, clearly benefitted from the ACT intervention, in terms of changes on outcome measures, and reported an increase in fusion and inflexibility pre- to post intervention, followed by a decrease in both variables, by 6-month follow-up. Changes on these ACT process measures with this patient group needs further investigation, as a similar increase in a related measure (the AAQI) has not been found in other ACT PD trials (e.g. Gratz and Gunderson, 2006).

The eventual increase in flexibility and defusion occurred during the same time period as the increases in valued action and improvements in Axis I

symptomology discussed earlier. Due to the study design and small sample size, it is not possible to test if there is a causal link between changes in the ACT processes and changes in outcome measures, or what the direction is of any such link. ACT theory suggests that these processes are implicated in psychopathology, and published outcome trials have shown related processes mediating ACT-related outcomes in (non-PD) treatment trials (e.g. Zettle et al., 2011.)

Finally, as with the previous study, the majority of the participants in this trial (90%) met diagnostic criteria for PDs from two clusters, with 40% having PD diagnoses across all three clusters. These results support the view that PD diagnostic categories do not represent completely separate disorders, and that there appears to be shared personality pathology across clusters.

8.4.2 *Methodological Limitations*

The majority of the limitations of this study, such as small sample size and lack of randomisation, are identical to those identified for the previous trial, and stem from design decisions appropriate to the initial testing of a new treatment for a potentially high risk patient group. These limitations are discussed in Chapter VII, Section 7.4.2. As with the previous study, the current sample did not include male participants, despite there being no gender difference in PD prevalence. This limitation means that generalising from the findings should be done with caution.

A number of important questions about possible causal relationships amongst study variables have been raised in this study, which could only be answered through significant changes in study design. Several limitations of the previous treatment protocol were addressed in this study, and this appeared to have a positive impact in terms of outcomes. However, the psychopathology outcomes were still relatively modest, and it may be that further fine-tuning of the protocol is required. Several recommendations for further research, addressing the current limitations, are made in Section 8.4.3.

8.4.3 *Implications and Future Research*

Although the number of participants in this trial was small, and so generalising from the findings should be done with this in mind, taking into

consideration the results of the first trial of the protocol as well, a number of possible recommendations for future ACT outcome research with this patient group suggest themselves.

8.4.3.1 *Participant Selection*

1. It seems likely that a period free from self-harm of longer than 6-months (possibly 12 months) prior to intervention would decrease the likelihood of self-harm during ACT.
2. Similarly, a period of more than 6-months free from substance misuse would be advisable.
3. A minimum of 6 months between the end of DBT and the start of the ACT intervention would be advisable.
4. Incidents of self-harm between DBT ending and ACT commencing may indicate the possibility of continued behavioural instability during the ACT intervention.
5. A participant's number of PD diagnoses did not appear to affect outcome, but specific diagnoses, (schizotypal and antisocial PD) might be associated with poor outcomes. Excluding people with these diagnoses might be advisable.

8.4.3.2 *ACT Protocol*

Based on changes in study measures and anecdotal reports, a number of aspects of the intervention may have been particularly helpful. These features of the protocol should be retained in any future version.

1. Connecting with personal values appears to have been motivating for many participants.
2. Repeatedly coaching participants in taking small behavioural steps in a valued life direction, appears to have been effective in increasing valued action.
3. A number of the participants who benefitted from the intervention appeared to bring about some change in their relationship with their private experiences such as thoughts, and cited both defusion and mindfulness practices as having been helpful in achieving this.

8.4.3.3 *ACT Protocol Changes*

1. In this study, many of the positive changes occurred in the 6-month follow-up period. This suggests that this patient group took a relatively long time to begin to make therapeutic gains, and therefore extending the intervention over a longer time period might optimise these improvements. Alternatively, it may be that extending the intervention is not required, but that the improvements might continue to develop beyond the current 6-month follow-up point, and that an extended follow-up period of 12 or 18 months is required.
2. The finding that participants initially became more fused and less flexible requires further investigation. If this is found to be detrimental for some participants, then further modifications to the protocol might be necessary, along the lines of further emphasising the flexible use of both avoidance/control and acceptance strategies.

8.4.3.4 *Measures*

1. All measures included in the study were completed by all participants without difficulty, suggesting feasibility of use with this patient group. As this was the first time some of the measures had been used with a PD sample, this research has established that they all could be considered for inclusion in future trials.
2. The VLQ was used in a simplified form and performed well with this patient group. There is no psychometric data available for this form of the measure, so a study examining its reliability and validity would be a useful development.
3. The LAS was developed for use in this study because the ACT therapists were concerned in Study 4 that anecdotal reports of increased positive patient activity had not been fully reflected in the psychometric data for that study. However, the 'amount of activity' variable of the LAS showed no change between any of the study time points, while group change on the VLQ valued action variable was significant. It seems likely that the LAS was not sensitive enough to detect these important changes, and its exclusion from future trials should be considered, unless it is significantly modified and tested psychometrically.

8.4.3.5 *Future Trial*

Although few significant results were found in this study, the effect size calculations suggested that an adequately powered study would have been likely to show several significant positive changes. The next step therefore should be an RCT, testing the protocol against a control condition of people with poor personality functioning who are receiving standard psychiatric care. The effect size calculations for the group changes in the 3 months prior to the start of the intervention suggest that participants were deteriorating. If this deterioration were to be observed in a control group, along with the improvements in functioning that would be predicted for the ACT condition, then large between-group differences, in favour of ACT, would be likely.

8.4.4 *Summary*

This study indicates that this ACT-based group intervention is a safe and acceptable treatment approach, post-DBT, for this potentially risky and difficult to treat group, and has the potential to help patients develop valued "lives worth living" (Linehan, 1993, p. 172). Taking into consideration the recommendations outlined above, the next step in testing the protocol would be to use an RCT design, with greater participant numbers and an extended follow-up period.

CHAPTER IX

General Discussion

As described in Chapter I, PDs are relatively common, chronic mental health problems, associated with risky behaviours and poor treatment outcomes. Although current diagnostic systems treat PDs as distinct disorders, there is evidence to suggest that there is personality pathology that cuts across PD diagnostic categories, and that PDs might be more helpfully be conceptualised dimensionally, in terms of poor personality functioning. There are some effective psychosocial interventions for PD, most notably DBT for BPD. DBT graduates however often continue to experience significant mental health problems, and report difficulties in engaging in life, and have been described as experiencing "quiet desperation" (Dimeff & Linehan, 2001, p.2). For other PD diagnoses, little treatment development research has been carried out and there are few evidence-based psychotherapy options.

There are two plausible reasons for considering ACT as a possible candidate for a post-DBT intervention for people with poor personality functioning across diagnostic categories. First, it was developed to impact universal psychological processes such as CF, and is therefore hypothesised to be effective across mental health diagnostic categories. Second, ACT emphasises experiencing rather than avoiding private experiences, and engaging in a valued life, and is therefore consistent with Linehan's (1993) views on what Stage II and III DBT interventions should address.

Based on these considerations, this thesis had two main aims: firstly, to investigate CF, an important yet under-researched component of the ACT model of psychopathology, and secondly, to develop and test a novel ACT intervention for a post-DBT sample with poor personality functioning. The thesis focused on phases I and II of the complex interventions development model outlined in Chapter III (Campbell et al., 2000). Initially, analogue studies were used to investigate CF, a crucial aspect of the model underpinning ACT. This included testing its relevance to personality functioning. This was followed by small-scale treatment-development studies, designed to pilot a new ACT-based group intervention, for a heterogeneous post-DBT PD sample.

This chapter provides a summary of the research findings, outlining the role of each study in the overall research programme. The thesis is also critically reviewed in terms of its main strengths and limitations. The implications of the study findings in relation to developing and refining ACT theory, and understanding and treating PDs, are also reviewed. Finally, recommendations for future research are made.

9.1 Main Findings

9.1.1 *Analogue Research*

ACT theory hypothesises that a number of linked psychological processes, including CF, underpin psychopathology. CF is a key component of the ACT psychopathology model, yet it remains relatively under-investigated, in large part due to the lack of a reliable and valid measure. Such a measure (the CFQ) has recently been developed, and Study 1 outlined its clinical validation. Based on a mixed, NHS, mental health sample (including people with PD diagnoses), and using CFA methodology, the CFQ was shown to have the same theory-consistent factor structure as had been identified with a non-clinical sample. Overall, it was found to be a reliable and valid self-report measure of CF, with the clinical sample. These findings allowed for the possibility of CF to be investigated with clinical and nonclinical samples.

Study 2 demonstrated one such application of the CFQ; testing the role played by CF in relation to personality functioning. The process was shown to fully mediate the relationships between known genetic (negative affectivity) and environmental (childhood trauma) risk factors for poor personality functioning, and actual personality functioning in adulthood. This model-testing study, although cross-sectional in design and therefore not able to demonstrate causality, represented the first evidence that CF is associated with poor personality functioning. Given that it is not possible to change an individual's history or genetics, it is important to identify mediating variables (CF in this case), to guide the development of interventions.

The final analogue study in this thesis was designed to use the CFQ in the development of a measure of the behavioural aspects of EA, another key ACT process. Avoidant behaviour is common amongst PD patients, and along with CF appears to interfere with healthy psychological functioning and engagement in life.

The study was based on a non-clinical, student sample, as a safe, first step. A prototype computer-based task was developed, which—based on the study findings—appeared to function as a behavioural measure of CF rather than EA. Although this result was unexpected, the availability of two measures of CF that utilise different methodologies, can only be advantageous in terms of increasing confidence CF as a construct, and the empirical evidence relating to CF.

The results from these analogue studies, along with existing research (e.g. Gratz and Gunderson, 2006; Clarke, et al., in prep) suggest that ACT, designed as it is, to impact universal psychological processes such as CF, might be of benefit to people with poor personality functioning. A novel ACT-based intervention for post-DBT patients with mixed PD presentations was therefore developed and initial testing of that protocol was carried out (see next section).

9.1.2 *Applied Research*

Previous outcome trials (e.g. Gratz & Gunderson, 2006), indicated that ACT might be beneficial in relation to PD, but its applicability to people with broad, transdiagnostic PD symptomology had not been tested. Given the complexity and risk histories of the target participants, a cautious approach to treatment development was taken.

Study 4 used an uncontrolled, pre-post design to test an ACT-based (with some DBT features) group intervention. Although the protocol was found to be feasible, and participants reported high levels of satisfaction, outcomes were mixed. There were no significant pre to post group differences, possibly due to the small sample size. RCI calculations indicated that the protocol had a positive impact on some primary outcome measures such as valued action, engagement in life, and some personality domains, and on some key process variables including CF. However, a minority of participants showed no improvements or deteriorated on some measures, particularly those assessing Axis I symptomology. Following feedback from participants and the therapists' experiences, several possible improvements to the protocol suggested themselves, and a second, small-scale trial was carried out to test the modified protocol.

Study 5, used the same design to test the extended and refined intervention. The results suggested that the changes to the protocol had had a positive impact. Encouraging results from the first trial, such as increases in valued activity, were also found in this trial, and were maintained or improved at 6-months follow-up. Additionally, effect size and RCI calculations indicated that the second version of the protocol had a more positive impact than the first, on psychiatric symptomology such as depression and PD diagnoses. Nonetheless, these latter improvements were relatively modest, and not all participants benefitted.

Changes in CF and psychological flexibility were consistent with each other; initial increases in fusion and inflexibility were followed by large improvements during the follow-up period, with the protocol being associated with a reduction in CF, overall. The improvements in these processes occurred at the same time as improvements in psychiatric symptomology.

Participants from both studies had highly complex personality presentations, with, on average, four PD diagnoses. The severity of Axis I symptomology in the cohorts matched that of *pre-treatment* participants in published CBT and DBT PD trials. For example, the mean score on the BDI-II (depression) across the two clinical studies in this thesis was 33.43. Davidson et al. (2006) reported a pre-intervention mean of 31.30 in their CBT condition, while McMain et al. reported a mean BDI-II score of 37.19 in their DBT condition. This is interesting given that the thesis study participants were DBT graduates, and DBT outcome research suggests that they would have experienced some improvements in Axis I functioning as a result of DBT. There is some evidence from Study 5 that the participants had deteriorated in terms of psychological functioning in the 3 months prior to starting ACT. This might suggest that DBT had improved Axis I symptomology but that those improvements had not been sustained. It is also possible that this group represents a particularly severe sub-group of DBT graduates who did not benefit from DBT in terms of Axis I disorders.

Both studies have limitations (discussed below), and should be viewed as the initial steps in the development of an effective treatment. Nonetheless, some provisional conclusions can be drawn from the findings, although generalising from these conclusions should be done with caution:

1. Based on the severity of pre-treatment symptoms, and verbally reported life difficulties, there is a need for a post-DBT psychosocial intervention for at least some DBT graduates.

2. ACT can be delivered to this patient group in a format that is feasible, acceptable and safe.

3. ACT can impact positively on engagement in life, valued action, and psychological functioning with this patient group.

4. Some of these improvements are modest, and there are indications that the protocol was not beneficial to all participants.

5. ACT-relevant processes such as CF appeared to be affected by the protocol, although not in a straightforward way, in the case of Study 5.

Taken together, the two clinical trials demonstrated the possibility of using an ACT-based group intervention to safely address post-DBT residual difficulties, for a particularly complex patient group. However, it is unclear whether the somewhat limited positive impact of the intervention might be improved by further refinement, or whether it indicates the limit of a non-specific, transdiagnostic approach, with further therapeutic gains only being possible with more focussed interventions. Alternatively, it may be the case that this particularly complex sub-group of people with PDs is especially difficult to treat, and these results, modest though they are, should be viewed as a validation of the approach taken.

9.1.4 *Main Strengths*

This programme of research had several strengths. I identified gaps in the current ACT model-testing literature and the PD treatment literature, and a novel, theory-consistent approach was taken to address these gaps in a coherent and systematic fashion. The programme of research that ensued was designed to fit within a well-established model for the development of complex psychosocial interventions. A multi-method approach was used, involving both analogue and applied research, to allow for a broad investigation of ACT and its relevance to personality functioning, across a range of samples. Diverse and advanced statistical methods supported this approach. The thesis has enhanced understanding of common yet significantly under-researched mental health difficulties, and involved the development of two measures

that have broad applicability, and a novel treatment approach for a patient group who are generally poorly served in terms of evidence-based interventions.

9.1.5 *Main Limitations*

A number of limitations must be weighed against the strengths reviewed above.

Analogue Studies Due to time and resource constraints, in Study 1, CFA was not preceded by EFA with a separate sample. Instead, the CFQ factor structure identified with non-clinical samples was used to suggest what factor structure should be modelled and tested with the clinical sample. This is the correct procedure when there are limited samples available, but ideally, two clinical samples would have been recruited. Similar constraints also meant that in Study 3, the computer-based task was not further developed to yield a measure of behavioural avoidance, as originally planned. This meant that although the study enhanced knowledge and resulted in the development of a behavioural measure of CF, some of the original aims of the study were not met. Finally, both for ethical and resource reasons, non-clinical samples were used in Studies 2 and 3 to investigate clinically-relevant phenomena.

Applied Studies Because Studies 4 and 5 involved a previously untested protocol and participants with significant risk histories, the ACT intervention was tested on very small samples, and no RCT was conducted. This means that it is unclear how well the findings from these studies will generalise, and without the random allocation of participants, a causal relationship between the intervention and the observed outcomes cannot be certain. These studies should therefore be viewed as initial steps on the path to developing an evidence-based intervention (see Section 9.2.5 for recommendations for next steps).

Also, in terms of limitations, as discussed in Section 9.1.2, the impact of the ACT protocol was somewhat limited, with some participants clearly not benefitting, suggesting that further development and testing is needed.

General Limitations As with the majority of related published research, all the thesis studies relied to a large extent on self-report measures. There are sound ethical and

economic justifications for this, but nonetheless, the use of such measures carries the risk of demand bias, participant fatigue, and inaccurate memory of experiences. Attempts were made to minimise these risks, such as allowing participants to complete questionnaires in private, and anonymously where possible. Additionally, Study 3, in yielding a behavioural measure of CF, will help improve this issue in future research.

Finally, in terms of limitations across studies, there were biases in all study samples in terms of gender, ethnic background, and in the case of Study 3, age.

9.2 Implications and Suggestions for Future Research

9.2.1 *CF, the CFQ, and the ACT Model*

The ACT model implicates a set of related processes in psychopathology, cutting across syndromal categories. The way in which these processes, including CF, relate to each other, is currently imprecisely articulated in the ACT literature, and it will be difficult to examine this issue empirically until there are valid measures of all relevant processes. The development and validation of the CFQ, the first psychometrically sound, general measure of CF, is therefore an important development for the ACT research community because it will support the examination of the role and impact of CF in many settings, with clinical and non-clinical populations. As with any psychometric measure, the CFQ itself should be subject to an on-going programme of testing and validation, including examining its performance with other physical and mental health populations, with samples from different ethnic backgrounds, and in a variety of settings.

There are several gaps in the ACT research literature in relation to CF that the CFQ could be used to address. For example, the role of CF as a possible mediator of therapeutic change in outcome trials is under-researched. Also, although the impact of stand-alone defusion practices has been tested (see Chapter II, Section 2.2.1.2), to date this work has relied on weak measures of CF. Replication of the findings from these studies using the CFQ, would be an important step in assessing the effect of defusion exercises.

Examination of the relationships within the ACT process model will require the development of good quality, distinct measures (such as the CFQ), for all of the processes involved. Related to this point, a new self-report measure of EA has been published since the studies in this thesis were planned (Gamez, Chmielewski, Kotov, Ruggero, & Watson, 2011). If it proves to perform well, it could contribute to the testing of the ACT model, as well as further validation of the CFQ and any behavioural measures of ACT processes such as the one developed in Study 3.

It will also be important, in terms of theory refinement, to examine the relationship between CF and other related constructs, such as decentering, and metacognitive awareness.

9.2.3 *Additional Mediators of Poor Personality Functioning*

Understanding mediating variables in relation to psychological difficulties is important because it can both increase theoretical understanding of those difficulties (through model-building), and guide treatment development. There is a paucity of modelling studies in relation to PD. Study 2 demonstrated that CF plays an important role in relation to poor personality functioning. Although CF fully meditated the relationship between risk factors and personality functioning, and accounted for 59% of the variance in the latter, other factors will also be relevant. Studies such as Cheavens et al., (2005) and Kingston et al., (2010), suggest that EA or a related construct such as thought suppression are likely candidates as additional mediators in relation to poor personality functioning. Also, given the hypothesised relationship between CF and EA whereby EA would not be necessary if fusion had not already occurred (e.g. Pistorello et al., 2000), the relationship between these constructs could be examined using a path model of personality functioning. Other constructs such as emotional regulation and self-compassion might also be expected to play a mediating role in relation to poor personality functioning, based on previous relevant research (e.g. Gratz & Tull, 2010; Gratz et al., 2006; Kuyken et al., 2010) and should be tested as part of a more comprehensive model of processes underpinning personality functioning.

9.2.4 *Behavioural Measures of Psychological Processes*

The potential risks associated with over-reliance on self-report measures have already been alluded to in this discussion. In clinically focussed ACT research in particular, behavioural measures that reflect hypothesised ACT processes are likely to increase the validity and utility of study findings. It can be a complex and time-consuming process to develop such measures to a point where they function as required, as seen in Study 3. The potential benefits however suggest that the development of both self-report and behavioural measures of key ACT processes should be pursued.

9.2.5 *Psychosocial Treatment of Poor Personality Functioning*

There are many gaps in our understanding of how to best treat PD, despite the pioneering work done by Linehan and others in the DBT research community, in particular. The studies in this thesis represent an attempt to address just one of those gaps: treating the continuing psychological difficulties of people with several PD diagnoses, post-DBT. The two linked treatment development studies have provided information about the impact of a novel ACT-based intervention for this patient group, as well as some indications of which patients might and might not benefit. In terms of the Rounsaville et al. (2001) treatment development stages, these studies have addressed Stage Ia. Stage Ib would involve an RCT with a larger sample, in order to assess the protocol's impact in a better powered and controlled study. Options for control conditions include a waiting list, treatment-as-usual, or an active control such as continued DBT, perhaps in the form of a DBT graduates' group.

This is of course not the only option for a post-DBT psychosocial intervention. Another possibility would be to include some features of ACT in a DBT Stage II intervention. Component studies testing specific aspects of ACT might aid clarification of which components of ACT to include. Based on clinical observation throughout the current clinical work, this is likely to include an emphasis on valued action, and the cultivation of a defused relationship to private experiences. A third option would be to develop and test a purely DBT Stage II intervention, as originally suggested by Linehan (1993). To date, no research testing such an intervention has been published.

The difficulty with the two latter possibilities is that DBT tends to be tailored to meet the needs of people with very specific PD presentations, usually BPD. Lynch and colleagues are in the process of testing a form of DBT for people with treatment resistant depression and co-morbid emotionally over-controlled PDs, cutting across several PD diagnoses (Lynch et al., 2011 – 2016). Their approach may well prove more effective for a heterogeneous PD patient group than the more general intervention tested in this programme of research (this would need testing empirically). However, there were several people in the current trials who simultaneously met the diagnostic criteria for both emotionally over- and under-controlled PDs, and it is unclear which currently available form of DBT would best meet their needs. If more than one of the above treatment options were pursued, the resulting protocols could be tested against each other in an RCT.

A significant advantage of an ACT-based protocol for a heterogeneous PD sample is that it is designed to impact universal psychological processes implicated in psychopathology. An alternative would be to develop a protocol based on a different approach, which is also thought to act through universal mechanisms. For example, Clarke et al. (2013) suggest that their CAT intervention for mixed PDs may have its impact through the therapeutic relationship and its impact on relationship with self and others. Based on their findings, this approach might have something to offer post-DBT patients with poor personality functioning across PD categories. However, the disadvantage of using CAT as a post-DBT therapy is that the two approaches are so different to each other (theoretically and practically) that the transition for patients from one to the other might prove difficult. ACT and DBT have sufficient shared theoretical underpinnings and focus of content that the transition is manageable.

Studies 4 and 5 confirmed the need for a post-DBT intervention. There are several approaches including, but not limited to, ACT, that might be of benefit to people with complex PD presentations, post-DBT. A great deal more treatment development research will be needed in order to ascertain the most effective approaches.

9.3 Concluding Remarks

This thesis was designed to enhance understanding and measurement of a specific aspect of the ACT model, CF, and to test the applicability of ACT, both

theoretically and clinically, to poor personality functioning. This integrated programme of research, although clearly just the initial steps towards a comprehensive ACT-based understanding and treatment of PD, has demonstrated that CF can be measured successfully with mental health patients including those with PDs, has measurable behavioural consequences, and is relevant to the development of personality problems. An ACT-based protocol has also been shown to have some positive effects for this highly complex patient group.

Levin at al. (2012) argue that it is not sufficient to develop and test the efficacy of psychosocial interventions; researchers should also be testing and refining relevant theory in parallel, with each of these strands of empirical work complementing and informing the other. Only is this way can psychotherapeutic interventions have a coherent theoretical basis, with empirically demonstrated mechanisms of change. This thesis represents an attempt to bring basic and applied science together in exactly this way. One possible advantage of a focus on process-orientated research alongside treatment development trials is that it might support an exploration of the commonalities (of action), between apparently differing therapeutic approaches, to the advantage of patients. Therapies such as ACT and DBT that have been hypothesised to have theoretical common ground (Hayes, 2004) might both be enhanced by joint empirical exploration.

The clinical trials in the thesis focused on an almost completely neglected sub-group of a relatively neglected patient group (people with PDs)—neglected both in terms of process and outcome research. They are a challenging group of people to include in research trials, but the suffering experienced by such patients and their loved-ones surely should motivate researchers to engage with this group. This thesis demonstrates that such work is possible.

APPENDICES

Appendix A: Summary of principal methodological characteristics and results for all cognitive behavioural therapies RCTs for PD, organised by intervention type

Author (s)	Participants	Intervention Group (IG)	Comparison Group (CG)	Test Points	Primary Measures	Results: Risk	Results: Axis I/ Other	Strengths/ Limitations
Behaviour Therapy								
Alden (1989)	Avoidant PD 20 - 40 years old All unmarried	Graded exposure (GE). 10 x weekly group sessions. N = 76 in total for across intervention and control groups - no N for specific conditions reported.	GE + social skills training. GE + social skills training + focus on intimacy. Wait list control. All active conditions matched for frequency of contact/ number of sessions.	Pre, post, 3-month follow-up.	SORT SRI SQ Ideographic measure of social targets. Self-monitoring of social activity. Interviewer behavioural rating.	No risk related information reported.	Significant improvements on all measures for all active conditions compared with w/list control, maintained at follow-up. Addition of skills training made no difference to outcomes.	Adherence-rating not independent. Comparison of 3 active, manualised conditions, and a w/list control. Short follow-up. Small N. No power calculation. Focused on PD rarely investigated. Data not analysed on ITT basis.

Author (s)	Participants	Intervention Group (IG)	Comparison Group (CG)	Test Points	Primary Measures	Results: Risk	Results: Axis I/ Other	Strengths/ Limitations
Cognitive Behaviour Therapy								
Davidson, Tyrer, Tata, Norrie, Palmer & Murray (2006)	BPD: 100% DSH or hospitalisation within previous 12 months	CBT + TAU. 30 x 1-to1 sessions over 1 year. N = 54	Non-Manualised, general UK NHS TAU. Not matched to test condition for frequency of contact/ number of sessions. N = 52.	Pre, post, 6 month, 12 month follow-up.	Suicidal acts, A & E visits, hospitalisations. ADSHI BDI-II STAI BSI IIP SFQ YSQ EuroQol WAI	Significant reduction in suicidal acts, A & E visits, hospitalisations for both conditions at follow-up. Significant between group difference for suicide acts, in favour of CBT.	Significant improvement on all measures for both conditions at 12 and 24 months. CBT outperformed TAU on symptom distress, state anxiety, dysfunctional beliefs.	Manualised and adherence rated treatment condition. Control condition neither manualised nor adherence rated. Controlled follow-up. Sufficiently powered.
Emmelkamp, et al., (2006)	Avoidant PD Excluded suicide risk.	Manualized CBT: 20 sessions over 6 months. N = 21	1. Manualized brief dynamic therapy: 20 sessions. N = 23 2. Waiting list. N = 18	Pre, post, 6 month follow-up.	SCID-II PDBQ LWASQ SPAI AS	Not applicable (high risk was an exclusion criteria).	Significant improvements on all measures for both conditions. CBT outperformed BDT and WLC	Manualised and adherence rated treatment and control condition, and WLC. Follow-up not including w/list control.

Author (s)	Participants	Intervention Group (IG)	Comparison Group (CG)	Test Points	Primary Measures	Results: Risk	Results: Axis I/ Other	Strengths/ Limitations
Emmelkamp et al., cont.							on majority of measures; maintained at 6-month follow-up.	Small N. No power calculation. Focused on PD rarely investigated. Data not analysed on ITT basis.
Evans, et al. (1999)	Cluster B personality "disturbance" + recent DSH.	Brief, CBT/DBT hybrid bibliotherapy + up to 6 sessions (MACT). N = 18.	General, non-manualized psychiatric TAU, Not matched to test condition for frequency of contact/ number of sessions. N = 16.	Pre, 6-months.	Time to next DSH episode. Median DSH per month. HADS SFQ Financial cost of follow-up per month.	Non-sig trend favouring MACT on time to next DSH episode and median DSH episodes.	MACT outperformed TAU on HADS only	Manualized treatment condition. Adherence not rated. No active control condition. Underpowered. Data not analysed on an ITT basis. Easy to implement, low-cost intervention.

Author (s)	Participants	Intervention Group (IG)	Comparison Group (CG)	Test Points	Primary Measures	Results: Risk	Results: Axis I/ Other	Strengths/ Limitations
Svartberg, Stiles & Seltzer (2004)	One or more Cluster C PD. Excluded all other PD diagnoses. Excluded parasuicidal behaviour.	CBT (40 weekly sessions). Conditions matched for frequency of contact/ number of sessions. N = 25.	Short term (40 weekly sessions) 1-to-1 dynamic psychotherapy (STDP). N = 25.	Pre, mid-therapy, post, 6, 12, and 24-month follow-up.	SCL-90 IIP MCMI-III	Not applicable (high risk was an exclusion criteria).	Significant improvements on all measures for both conditions. No significant between group effects. Some trends in favour of STDP.	Adherence independently rated. Comparison of 2 active, manualised interventions. Controlled, long follow-up. Under powered. Focused on PDs rarely investigated.
Blum, St. John, Pfohl, Stuart, McCormick, Allen et al. (2008)	BPD: not required to have recent parasuicidal acts.	Systems Training for Emotional Predictability and Problem Solving (STEPPS) 20 x weekly group sessions. + TAU. N = 65.	General, non-manualised TAU, could include 1-to-1 therapy. Not matched to test condition for frequency of contact/ number of sessions. N = 59.	Pre, post, 1 year follow-up.	ZRSBPD BESOT PANAS CGI GAS BDI SCL-90 BIS SAS	Significant reduction in suicidal acts,self-harm, hospitalisations for both conditions, maintained at follow-up. No significant between group differences.	Significant improvement on all measures for both conditions, maintained at follow-up. STEPPS outperformed TAU on ZRSBPD, PANAS, CGI, GAS, BDI, BIS	Manualised/ adherence rated treatment condition. Non-manualised TAU control. Controlled follow-up/ Sufficiently powered. High attrition in STEPPS condition.

Author (s)	Participants	Intervention Group (IG)	Comparison Group (CG)	Test Points	Primary Measures	Results: Risk	Results: Axis I/ Other	Strengths/ Limitations
Dialectical Behaviour Therapy								
Linehan, Armstrong, Suarez, Allmon, & Heard (1991)	BPD. Female. Parasuicidal act within previous 8 weeks.	DBT. Full programme. 12 months. N = 24.	Non-manualised, standard TAU. Not matched to test condition for frequency of contact/ number of sessions. N = 22.	Pre, 4, 8, 12 months.	PHI THI SSI BDI BHS RLISCS	Significant between group differences in favour of DBT on number of parasuicidal acts, medical risk associated with those acts, psychiatric inpatient days.	Significant improvement on all other measures for both conditions, No significant group differences.	No therapist adherence reported. Manualised treatment condition. Non-manualised TAU control. No follow-up reported. Underpowered. Low attrition for DBT. Data not analysed on ITT basis.
Linehan, Schmidt, Dimeff, Craft, Kanter, & Comtois (1999)	BPD + substance use disorder. Not required to have recent parasuicidal acts.	DBT. Full programme + replacement medications. 12 months. N = 12.	Non-manualised, standard TAU. Not matched to test condition for frequency of contact/ number of sessions. N = 15.	Pre, 4, 8, 12, 16 months.	Structured clinical interview re drug abuse. Urinalysis. THI PHI SHI	No between group difference on parasuicidal acts.	Significant group differences favouring DBT on drug abuse, global + social adjustment.	No adherence reported. Manualised treatment condition. Non-manualised control.

Author (s)	Participants	Intervention Group (IG)	Comparison Group (CG)	Test Points	Primary Measures	Results: Risk	Results: Axis I/ Other	Strengths/ Limitations
Linehan et al., cont.	Female				GSA GAS SAEI		No between group differences on other measures.	Short, controlled follow-up. Underpowered.
Turner (2000)	BPD. Recent history of suicide attempt.	DBT-oriented therapy including psychodynamic techniques. Skills training within 1-to-1 sessions, not as separate groups. 12 months. N = 12.	Manualised client-centred therapy. Matched to test condition for frequency of contact/number of sessions. N = 12.	Pre, 6, 12 months.	HRSD BPRS TBR BDI BAI BSSI	Significant improvement for both conditions on parasuicidal behaviours and hospitalisation. Significant between group differences in favour of DBT.	Significant improvement for both conditions on all other measures. Significant between group differences in favour of DBT on depression, global functioning, impulsivity and anger.	Adherence not reported. Manualised, active control. No follow-up reported. Underpowered. Naturalistic setting. DBT condition was not based on standard DBT.
Koons, Robins, Tweed, Lynch, Gonzalez, Bishop et al. (2001)	BPD. Female US veterans. 40% parasuicidal act in previous 6 months.	DBT. Full prgramme. 6-months. N = 13.	Non-manualised TAU. 1 x weekly 1-to-1 non-specified psychotherapy + psychoeducational groups.	Pre, 3, 6 months.	PHI BSSI BHS BDI HAM-D HARS SAES	Significant improvement for both conditions on parasuicidal behaviours. Significant	Significant improvement for both conditions on all other measures except anxiety	Manualised and adherence rated treatment condition. Non-

Author (s)	Participants	Intervention Group (IG)	Comparison Group (CG)	Test Points	Primary Measures	Results: Risk	Results: Axis I/ Other	Strengths/ Limitations
Koons, Robins, Tweed, Lynch, Gonzalez, Bishop et al. (2001) cont.	Antisocial PD excluded.		Partly matched to DBT for frequency of contact/number of sessions. N = 12.		DES	between group differences in favour of DBT.	(neither group changed) and suicidal ideation and hopelessness (control did not change). Significant between group differences in favour of DBT on suicidal ideation, hopelessness, depression, and anger expression.	manualised TAU control. Underpowered Data not analysed on ITT basis.
Linehan, Dimeff, Reynolds, Comtois, Welsh, Heagerty et al. (2002)	BPD + opiate dependance. Not required to have recent parasuicidal acts. Female.	DBT. Full prgramme + replacement medications. 12 months. N = 11.	Comprehensive Validation Therapy + 12-step programme + replacement medications. Partly matched to DBT for frequency of contact/number of sessions. N = 12.	Pre, 4, 8, 12, 16 months.	Urinalysis. PHI SHI GAF BSI	Significant reduction in suicidal acts, self-harm, hospitalisations for both conditions, maintained at follow-up. No significant between group differences.	Significant improvement in all outcome measures for both groups. DBT: better maintenance of reduction in drug use. No other between group differences.	Adherence not independently rated. Manualised conditions. Short, controlled follow-up. Underpowered Higher attrition in DBT (36%).

Author (s)	Participants	Intervention Group (IG)	Comparison Group (CG)	Test Points	Primary Measures	Results: Risk	Results: Axis I/ Other	Strengths/ Limitations
Van den Bosch, Schippers, Verheul, & van den Brink (2002); Verheul, van den Bosch, Koeter, de Ridder, Stijnen, & van den Brink (2003).	BPD +/- substance abuse. Not required to have recent parasuicidal acts. Female.	DBT. Full prgramme. 12 months. N = 31.	Non-manualised TAU. Not matched to test condition for frequency of contact/ number of sessions. N = 33	Pre, 11, 22, 33, 44, 52 weeks, 6-month follow-up.	Treatment retention. BPDSI LPC	Significant reduction in self-harm for both conditions. Significant between group difference in favour of DBT.	No significant improvement for either condition on substance misuse.	Manualised treatment condition. Adherence not reported. Non-manualised control. Controlled follow-up. Power calculation not reported. High attrition rate (77%) in control. Naturalistic setting.
Linehan et al. (2006); Harned, et al., (2008)	BPD + recent parasuicidal acts. Female.	DBT. Full programme. 12 months. N = 52.	Non-manualised Community Treatment by Experts. Not matched to test condition for frequency of contact/ number of sessions. N = 51.	Pre, 4 monthly, to 12-month follow-up.	SASII SBQ RLI THI HRSD LIFE	Significant reduction in suicide attempts, self-harm and hospitalisation for both conditions. Significant between group differences in	Significant improvement on depression, hopelessness, , suicidal ideation, Axis-I disorders, and reasons for living for both groups. No between	Manualised/ adherence rated treatment condition. Non-manualised control. Controlled follow-up. Sufficiently powered.

Author (s)	Participants	Intervention Group (IG)	Comparison Group (CG)	Test Points	Primary Measures	Results: Risk	Results: Axis I/ Other	Strengths/ Limitations
						favour of DBT on suicide attempts and hospitalisation. No significant between group differences on self-harm.	group differences on these variables except for substance dependence (in favour of DBT).	
McMain, Links, Gnam, Guimond, Cardish, Korman et al. (2009)	BPD + parasuicidal act within past 5 years.	DBT. Full programme. 12 months. N = 90.	Non-manualised general psychiatric management TAU. Not matched to test condition for frequency of contact/ number of sessions. N = 90.	Pre, 4-monthly. No follow-up reported.	SASII ZRSBPD SCL-90 STAEI BDI IIP THI RETFTQ	Significant reduction in suicide attempts, self-harm and hospitalisation for both conditions. No significant between group differences on these variables.	Significant reduction in BPD symptoms, depression, interpersonal functioning, symptom distress and anger for both conditions. No significant between group differences on these variables.	Manualised and adherence rated treatment condition. Non-manualised control. No follow-up reported. Sufficiently powered.

Measures abbreviations:

SORT: Social Reticence Inventory
SQ: Shyness questionnaire
BDI: Beck Depression Inventory
BSI: Brief Symptom Inventory
SFQ: Social Functioning Questionnaire
WAI: Working Alliance Inventory
CGI: Clinical Global Impression
SAS: Social Adjustment Scale
BESOT: Borderline Evaluation of Severity Over Time
PSI: Parasuicide History Interview
SSI: Scale for Suicidal Ideators
RLISCS: Reasons for Living Inventory, Survival and Coping Scale
SAEI: State-Trait Anger Expression Inventory
BSI: Borderline Syndrome Index
BAI: Beck Anxiety Inventory
HAM-D: Hamilton Depression Rating Scale
BPDSI: BPD Severity Index
SASII: Suicide Attempt Self-Injury Interview
LIFE: Longitudinal Interval Follow-up Evaluation
EQ-5D: EuroQol Quality of Life
Revised.Interview
BPRS: Brief Psychiatric Rating Scale
SCID-II: Structured Clinical Interview for DSM-IV Axis II Disorders
LWASQ: Lehrer Woolfolk Anxiety Symptoms Questionnaire
AS: Avoidance Scale
SRI: Self-Report Inventory
ADSHI: Acts of Deliberate Self-Harm Inventory
STAI: Spielberger State Trait Anxiety Inventory
IIP: Inventory of Interpersonal Problems
YSQ: Young Schema Questionnaire
ZRSBPD: Zanarini Rating Scale for BPD
GAS: Global Assessment Scale
SCL-90: Symptom Checklist-90
PANAS: Positive and Negative Affect Schedule
THI: Treatment History Interview
BHS: Beck Hopelessness Scale
SHI: Social History Interview
HRSD: Hamilton Rating Scale for Depression
TBR: Target Behaviour Ratings
BSSI: Beck Scale for Suicidal Ideation
HARS: Hamilton Anxiety Rating Scale
LPC: Lifetime Parasuicidal Count
SBQ: Suicidal Behaviours Questionnaire
RFETFTQ: Reasons for Early Termination from Treatment Questionnaire
DIB-R: Diagnostic Interview for Borderline Personality Disorders-

for DSM-IV Axis II Disorders.
PDBQ: Personality Disorder Belief Questionnaire
SPAI: Social Phobia Anxiety Inventory

Summary of principal methodological characteristics and results for all psychodynamic and interpersonal RCTs for PD, organised by intervention type

Author (s)	Participants	Intervention Group (IG)	Comparison Group (CG)	Test Points	Primary Measures	Results: Risk	Results: Axis I/ Other	Strengths/ Limitations
Schema Therapy								
Giesen-Bloo, van Dyck, Spinhoven, van Tilburg, Dirksen, van Asselt, et al. (2006)	BPD: 50% DSH within previous 3 months. Antisocial PD excluded	Schema Therapy (ST) 2 x weekly individual sessions for 3 years N = 44	Transference Focused Therapy (TFT) 2 x weekly individual sessions for 3 years. Matched with test condition for frequency of contact/ number of sessions. N = 42	3-monthly for 3 years	BPDSI-IV EuroQol WHOQOL Composite psycho- and personality pathology measure	Significant reduction in para-suicidal acts for both conditions. ST superior to TFT	Significant improvement on all measures for both conditions at 1, 2, 3 years. ST superior to TFT on all variables except QoL	Adherence & competence rated. Comparison of 2 active, well-described interventions. Controlled follow-up. Underpowered. Hospitalisation not measured.
Farrell, Shaw, & Webber (2009)	BPD: 100% DSH within previous 24 months	Schema Therapy + TAU 30 x group sessions over 8 months N = 16	Non-manualised TAU. Not matched with test condition for frequency of contact/ number of	Pre, post, 6-month follow-up.	BSI SCL-90 DIB-R GAFS	Significant reduction in self-injurious and impulsive behaviours for ST.	Significant improvement on all measures for ST, but not control condition Significant group differences on all measures	Adherence not assessed independently. Poor quality TAU Controlled follow-up. Small N. No power

Author (s)	Participants	Intervention Group (IG)	Comparison Group (CG)	Test Points	Primary Measures	Results: Risk	Results: Axis I/ Other	Strengths/ Limitations
Farrell et al. cont.			sessions. N = 12					calculation. 100% retention rate in ST
CAT								
Chenen, Jackson, McCutcheon, Jovel, Dudgeon, Yeun et al. (2008)	Adolescents with 2 - 9 BPD diagnostic criteria	CAT Maximum of 24 weekly, 1-to-1 sessions N = 44	Manualised 'good clinical care' (GCC). Matched to test condition for frequency of contact/ number of sessions. N = 42	Pre, post, 6-monthly for 2 years follow-up.	SCID-II YSR SOFAS Number of parasuicudal episodes	Significant reduction in self-injurious and impulsive behaviours for both conditions. No significant group differences.	Significant improvement on all measures for both conditions, maintained at 24-months. No significant group differences.	Adherence & competence rated. Good quality control, though not specifically designed for PD. Controlled follow-up. Underpowered.
Clarke et al. (in prep)	At least one PD diagnosis. Exclusion of current parasuicidal behaviour.	CAT 24 sessions 1-to-1 over 10 months. N = 50.	Non-manualised, standard NHS care TAU. Not matched to test condition for frequency of contact/ number of sessions. N = 49.	Pre, post, and uncontrolled follow-up at 9 and 18 months.	CORE IIP DISQ SCL-90 DES PSQ SSS	N/A (high risk was an exclusion criteria).	Significant group differences favouring CAT on SCID-II, IIP, CORE, PSQ, DIS-Q. No group differences on other measures.	Adherence & competence rated. TAU not specifically for PD. Uncontrolled follow-up. Underpowered. High attrition at follow-up.

Author (s)	Participants	Intervention Group (IG)	Comparison Group (CG)	Test Points	Primary Measures	Results: Risk	Results: Axis I/ Other	Strengths/ Limitations
Clarke et al. cont.								Naturalistic setting. Included PDs not often investigated.
Psychodynamic								
Munroe-Blum & Marziali (1995)	BPD Parasuicidal behaviour not excluded > 1/3 made a suicide attempt in 6 months prior to study.	Interpersonal group. 25 x weekly sessions. 5 x biweekly sessions. N = 38.	Individual psychodynamic psychotherapy. 1 or 2 x weekly. Open ended. Greater number of sessions and frequency of sessions than test condition. N = 41.	Pre, 6-monthly for 24 months.	OBI SAS BDI HSCL-90	Significant reduction in risk behaviours for both groups, maintained at 24 months. No between group differences.	Significant reduction in depression, general psychopathology and improvement in social functioning for both groups, maintained at 24 months. No between group differences.	Adherence not assessed independently. Control condition: experienced therapists/non manualised, homogenous intervention. Controlled follow-up. High dropout rates post-randomisation.
Bateman & Fonagy (1999; 2001)	BPD DSH not an inclusion	Partial hospitalisation (PH).	TAU Standard NHS psychiatric	3-monthly for 36 months.	SSHI SCL-90 BDI	Significant reduction in self-harm,	Significant improvement on all other	Adherence rating based on self-report.

Author (s)	Participants	Intervention Group (IG)	Comparison Group (CG)	Test Points	Measures	Results: Risk	Results: Axis I/ Other	Strengths/ Limitations
Bateman & Fonagy cont.	criteria. Median of 9 episodes in 6 months prior to study. 95% suicide attempt in 6 months prior to study.	Psychodynamic group & 1-to-1 Total of 6 x week contact minimum max. 18 months N = 19.	care. Not matched to test condition for frequency of contact/ number of sessions. N = 19.		STAI SAS IIP	suicide attempts and days in hospital for PH, but not for control.	measures for PH but not control.	Less intensive, non-manualised control. PH group had therapy during follow-up. Underpowered. Naturalistic setting. Compromised randomisation. Included high risk patients.
Bateman & Fonagy (2009)	BPD. Parasuicidal act within previous 6 months.	MBT: Psychodynamic group & 1-to-1 over 18 months. N = 71.	Structured clinical management (SCM). N = 63.	Pre, 6, 12, 8 months.	Suicide attempts Life-threatening DSH Hospital admission SCL-90 BDI SAS GAF IIP	Significant reduction in suicide attempts, DSH and days in hospital for both conditions, with MBT outperforming SCM on all measures at 18 months.	Significant improvement on all other measures for both conditions. MBT outperformed SCM on the GAF, IIP, SCL-90, SAS.	Manualised and adherence rated treatment and active condition condition. No follow-up reported. Sufficiently powered. Included high risk patients.

Author (s)	Participants	Intervention Group (IG)	Comparison Group (CG)	Test Points	Measures	Results: Risk	Results: Axis I/ Other	Strengths/ Limitations
Svartberg, Stiles & Seltzer (2004)	One or more Cluster C PD. Excluded all other PD diagnoses. Excluded parasuicidal behaviour.	Short term (40 weekly sessions) 1-to-1 dynamic psychotherapy (STDP). N = 25.	CBT (40 weekly sessions). Conditions matched for frequency of contact/ number of sessions. N = 25.	Pre, mid-therapy, post, 6, 12, and 24-month follow-up.	SCL-90 IIP MCMI-III	Not applicable (high risk was an exclusion criteria).	Significant improvements on all measures for both conditions. No significant between group effects. Some trends in favour of STDP.	Adherence independently rated. Comparison of 2 active, manualised interventions. Controlled, long follow-up. Under powered. Focused on PDs rarely investigated.
Vinnars, Barber, Noren, Gallop, & Weinryb (2005)	At least one PD diagnosis. High risk not excluded.	Supportive-expressive dynamic therapy (SEDT). 40 x 1-to-1 sessions. N = 80.	Non-manualised community-based psychodynamic therapy (CPT). Open number of sessions. Not matched to test condition for frequency of contact/ number of sessions.	Pre, post SEDT, 1 year follow-up.	SCID-II SCL-90 GAFS	Not reported.	Significant improvements on all measures for both conditions. No significant between group effects.	Manualised and adherence rated treatment condition. Control condition neither manualised nor adherence rated. Controlled follow-up. Sufficiently powered.

Author (s)	Participants	Intervention Group (IG)	Comparison Group (CG)	Test Points	Measures	Results: Risk	Results: Axis I/ Other	Strengths/ Limitations
Vinnars et al., cont.			N = 76.					Naturalistic setting. Included a range of PD diagnoses.
Muran, Samstag, Safran & Winston (2005)	At least one Cluster C or NOS PD diagnosis. Excluded Cluster A and B PD.	Alliance focused brief relational therapy (BRT). Manualised 30 x weekly 1-to-1 sessions. N = 33.	CBT. Manualised 30 x weekly 1-to-1 sessions. N = 29. Short-term dynamic therapy Manualised 30 x weekly 1-to-1 sessions. N = 22. All conditions matched for frequency and number of sessions	Pre, post, 6-month follow-up.	SCL-90 TC GAS IIP WPI	Not reported.	Significant improvements on all measures for both conditions. No significant between group effects.	All conditions adherence-rated. Comparison of 3 active, manualised conditions. Controlled follow-up. Underpowered. Naturalistic setting. Focused on PDs rarely investigated. Compromised randomisation process. Data not analysed on ITT basis.

Author (s)	Participants	Intervention Group (IG)	Comparison Group (CG)	Test Points	Measures	Results: Risk	Results: Axis I/ Other	Strengths/ Limitations
Clarkin, Levy, Lenzenweger & Kernberg (2007)	BPD. Other PDs not reported.	Transference-focused psychotherapy (TFP). Manualised, 1 year 2 x weekly 1-to-1 sessions. N = 30.	DBT: Manualised, 1 year full programme. N = 30. Supportive psychotherapy: Manualised, 1 year, 1 x weekly 1-to-1 sessions. N = 30. Conditions not matched for frequency/ number of sessions.	Pre, 4, 8, 12 months. No follow-up data reported.	OAS-M AIAQ BAI-II BSI BDI GAFS SAS	Significant reduction in suicidality for TFP and DBT only.	Significant improvements in depression, anxiety, global functioning, social adjustment for all 3 conditions. Significant improvements in anger, impulsivity for TFP and SP only.	All conditions adherence-rated. Comparison of 3 active, manualised, conditions. Controlled follow-up. Underpowered. Included males in sample.
Gregory, Chlebowski, Kang, Remen, Soderberg, Stepkovitch et al. (2008)	BPD with co-morbid alcohol addiction or dependence. Other PDs, including Antisocial PD not excluded.	Dynamic Deconstructive Psychotherapy (DDP). Manualised, 12 - 18 months 1 x weekly 1-to-1 sessions. N = 15.	TAU. Non-manualised, general psychiatric including referral to alcohol rehabilitation centre. 12-months.	Pre, 3, 6, 9, 12 months. No follow-up data reported.	LPC ASI THI BDI DES SPS BEST	Significant reduction in parasuicidal behaviour at 12 months. No between group difference.	Significant improvement on alcohol misuse, institautional care, BPD symptomology, depression, dissociation & perceived social support at 12 months for DDP.	Adherence not rated. Poor quality, non-manualised control. No follow-up data reported. Underpowered. Naturalistic setting.

Author (s)	Participants	Intervention Group (IG)	Comparison Group (CG)	Test Points	Measures	Results: Risk	Results: Axis I/ Other	Strengths/ Limitations
			Not matched to test condition for frequency and number of sessions. N = 15.				Significant between group differences on BPD symptomology, depression and perceived social support.	Included participants with severe problems.
Doering, Horz, Rentrop, Fischer-Kern, Schuster, Benecke et al. (2010)	BPD. Antisocial PD excluded.	Transference-Focused Psychotherapy (TFP). Manualised 1 year 2 x weekly 1-to-1 sessions. N = 52.	TAU. 1-to-1 psychotherapy by experienced community therapists. Not matched to test condition for frequency and number of sessions. N = 52.	Pre, post. No follow-up data reported.	CISSB SCID-I SCID-II GAFS BDI STAI CRTHI STIPO	Significant reduction in suicide attempts for TFP. Between group effect in favour of TFP. No significant reduction in self-harm for either condition.	Significant group effect in favour of TFP on BPD symptomology, psychosocial functioning, personality organisation. Both conditions showed significant improvement on depression, anxiety, with no group differences.	TFP adherence rated. Non-manualised control condition. No follow-up data reported. Adequately powered. High dropout rates in both conditions. Naturalistic setting.

Measures abbreviations:

BPDSI-IV: Borderline Personality Disorder Severity Index, fourth edition.
EuroQol: EuroQol quality of life thermometer measure.
WHOQOL: World Health Organization Quality of Life measure.
SCL-90: Symptom Checklist-90
Revised.
GAFS: Global Assessment of Function Scale.
SCID-I: Structured Clinical Interview for DSM-IV Axis I Disorders.
SOFAS: Social and Occupational Functioning Assessment Scale
CORE: Clinical Outcomes in Routine Evaluation
DISQ: Dissociative Questionnaire
PSQ: Personality Structure Questionnaire
DIBP: Diagnostic Interview for Borderline Patients
BDI: Beck Depression Inventory
SAS: Social Adjustment Scale
AIAQ: Anger, Irritability and Assault Questionnaire
BSI: Brief Symptom Inventory
ASI: Addiction Severity Index
SPS: Social Provisions Scale
CISSB: Cornell Interview for Suicidal and Self-Harming Behaviour
STIPO: Structured Interview for Personality Organisation
HSCL-90: Hopkins Symptom Checklist
BSI: Borderline Syndrome Index
DIB-R: Diagnostic Interview for Borderline Personality Disorders-

SCID-II: Structured Clinical Interview for DSM-IV Axis II Disorders.
YSR: Youth Self-Report questionnaire
MCMI-III: Millon Clinical Multiaxial Inventory
IIP: Inventory of Interpersonal Problems
DES: Dissociative Experiences Scale
SSS: Service Satisfaction Scale
SSHI: Suicide and Self-Harm Inventory
STAI: Spielberger State Trait Anxiety Inventory
OAS-M: Overt Aggression Scale-Modified
BIS-II: Barratt Impulsivity Scale-II
LPC: Lifetime Parasuicidal Count
THI: Treatment History Interview
BEST: Borderline Evaluation of Severity over Time
CRTHI: Cornell Revised Treatment History Inventory
OBI: Objective Behaviors Index

Appendix B: Cognitive Fusion Questionnaire

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

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

1. My thoughts cause me distress or emotional pain	1	2	3	4	5	6	7
2. I get so caught up in my thoughts that I am unable to do the things that I most want to do	1	2	3	4	5	6	7
3. Even when I am having distressing thoughts, I know that they may become less important eventually	1	2	3	4	5	6	7
4. I over-analyse situations to the point where it's unhelpful to me	1	2	3	4	5	6	7
5. I struggle with my thoughts	1	2	3	4	5	6	7
6. Even when I'm having upsetting thoughts, I can see that those thoughts may not be literally true	1	2	3	4	5	6	7
7. I get upset with myself for having certain thoughts	1	2	3	4	5	6	7
8. I need to control the thoughts that come into my head	1	2	3	4	5	6	7
9. I find it easy to view my thoughts from a different perspective	1	2	3	4	5	6	7
10. I tend to get very entangled in my thoughts	1	2	3	4	5	6	7
11. I tend to react very strongly to my thoughts	1	2	3	4	5	6	7
12. Its possible for me to have negative thoughts about myself and still know that I am an OK person	1	2	3	4	5	6	7
13. It's such a struggle to let go of upsetting thoughts even when I know that letting go would be helpful	1	2	3	4	5	6	7

Thank you for completing this questionnaire

Appendix C: Brief Demographics Questionnaire

Please complete this questionnaire by ticking the appropriate options. All information you supply will remain confidential and will be stored anonymously.

1. Are you male or female? Male _____ Female _____

2. How old are you? _____ years old

3. What is your ethnic origin? White _____ Asian _____

   Black _____ Mixed _____

   Other ethnic group _____

4. Are you currently having psychological treatment (e.g. counselling, psychotherapy)? Yes _____ No _____

THANK YOU FOR COMPLETING THIS QUESTIONNAIRE

Appendix D: Study 1 Research Ethics Committee Approval Letter

National Research Ethics Service
SOUTHAMPTON & SOUTH WEST HAMPSHIRE
RESEARCH ETHICS COMMITTEE (A)
1ST Floor, Regents Park Surgery
Park Street, Shirley
Southampton
Hampshire
SO16 4RJ

Tel: 023 8036 2466
023 8036 3462
Fax: 023 8036 4110

Email: scsha.SWHRECA@nhs.net

CM/sta

15 June 2009

Professor Susan E Clarke
Head of Intensive Psychological Therapies Service (IPTS),
Consultant Clinical Psychologist, Dorset Health Care NHS Foundation Trust
Dorset Healthcare NHS Foundation Trust
IPTS
51a Layton Road
Poole
BH12 2BJ

17 JUN 2009

Dear Professor Clarke

Study Title: **Validation of a self-report measure of cognitive fusion**
REC reference number: **09/H0502/78**
Protocol number: **1**

The Research Ethics Committee reviewed the above application at the meeting held on 09 June 2009. Thank you for attending to discuss the study.

Ethical opinion

The members of the Committee present gave a favourable ethical opinion of the above research on the basis described in the application form, protocol and supporting documentation, subject to the conditions specified below.

Ethical review of research sites

The favourable opinion applies to all NHS sites taking part in the study, subject to management permission being obtained from the NHS/HSC R&D office prior to the start of the study (see "Conditions of the favourable opinion" below).

Conditions of the favourable opinion

The favourable opinion is subject to the following conditions being met prior to the start of the study.

Management permission or approval must be obtained from each host organisation prior to the start of the study at the site concerned.

For NHS research sites only, management permission for research ("R&D approval") should be obtained from the relevant care organisation(s) in accordance with NHS research governance arrangements. Guidance on applying for NHS permission for research is available in the Integrated Research Application System or at http://www.rdforum.nhs.uk. Where the only involvement of the NHS organisation is as a Participant Identification Centre, management permission for research is not required but the R&D office should be notified of the study. Guidance should be sought from the R&D office where necessary.

Sponsors are not required to notify the Committee of approvals from host organisations.

Further to discussions during the meeting regarding the separation of sending out the

Appendix E: Study 1 consent form, participant information sheet and debrief sheet

CONSENT FORM

Title of the project: *Validation of a self-report measure of cognitive fusion.*

Name of Researcher: Prof. Sue Clarke

Please initial each box ↓

1.	I confirm that I have read and understood the information sheet dated 11/05/2009 (version 1) for the above study. I have had the opportunity to consider the information, ask questions by phone and have had these answered satisfactorily	
2.	I understand that my participation is voluntary and that I am free to withdraw at any time without my medical care or legal rights being affected	
3.	I understand that data collected throughout this study may be looked at by members of Prof. Sue Clark's research team. I give permission for these individuals to have access to the data collected from this study, to store and to process it.	
4.	I agree that in order for me to participate in the study, members of Prof. Susan Clarke's research team may have access to my contact details, which will be stored securely on an NHS, password-protected computer or locked in an NHS filing cabinet.	
5.	I agree for my GP to be informed about my participation in this study.	
6.	I agree that data obtained from this trial can be used, in anonymous form, for publication.	

_____	_____	_____
Participant name	Date	Signature
_____	_____	_____
Researcher name	Date	Signature

Participant information sheet

Validation of a self-report measure of cognitive fusion.

You are being invited to take part in a research study. Before you decide whether you would like to take part, you need to understand why the research is being done and what it would involve. Please take time to read the following information carefully. Talk to others about the study if you so wish.

- Part 1 tells you the purpose of this study and what will happen to you if you take part.
- Part 2 gives you more detailed information about the conduct of the study.

If there is anything that is not clear, or if you would like more information, please do not hesitate to contact Helen Bolderston or Sue Clarke on 01202 584120.

Part 1

What is the purpose of this study?
This study is designed to help develop a new questionnaire. The questionnaire measures a psychological process called 'cognitive fusion'. Cognitive fusion has been defined as: 'the process by which thoughts about an event become merged with the actual event', so that people experience, for example, thoughts about a possible loss as if the loss had actually occurred, feeling sadness and so on.

This process, experienced by everyone to some extent, is thought to have a role in the development and maintenance of a range of psychological problems, so it is important to be able to examine, measure and understand cognitive fusion. However, currently there is no good quality questionnaire available to measure cognitive fusion, hence the current study. This study has been designed by senior clinicians and researchers (Prof. Susan Clarke; Helen Bolderston).

Why have I been invited?
We are recruiting participants who are currently receiving care from local community mental health teams (CMHTs), or who are involved in assessment or psychotherapy at the Intensive Psychological Therapies Service (IPTS). We would therefore like to invite you to participate in the study. We aim to recruit a maximum of 200 participants.

Do I have to take part?
It is up to you whether you take part or not. If you do decide to take part, you will be asked to sign a consent form.

Having signed the consent form, you are still free to withdraw at any time and without giving a reason. A decision to withdraw, or a decision not to take part, will not affect the standard of care you receive, in any way.

What will happen to me if I take part; what will I have to do?

If you decide to take part, you will need to fill out and sign both copies of the enclosed consent form, keeping one copy for your own records and returning the other in the envelope provided. You will then be sent the questionnaires to fill out, which should take between 40 and 45 minutes in total. These are for you to complete in your own time, but it is important, once you begin filling out the questionnaires, that you complete them all within five days.

Once you have completed the questionnaires, they should be posted to the research team, in the stamped, addressed envelope provided. This would probably mark the end of your participation in the study. However, a small number of participants will be contacted once more, three weeks after returning this set of questionnaires. They will be asked to complete the questionnaire under development but none of the other questionnaires. This will take approximately five minutes.

If you are willing to participate, but would like help filling out the questionnaires, please contact Helen Bolderston or Sue Clarke on 01202 584120, as we would be happy to arrange this for you.

What are the questionnaires like?
There are eight questionnaires in total, which vary in length. There is a brief demographics questionnaire which asks about such things as your age and ethnic background. There are then seven further questionnaires, focussing on a range of psychological processes and problems.

Reactions to the questionnaires
All of the questionnaires except the one under development and the demographic questionnaire are widely used in research and clinical settings, and it is unlikely that you will experience distress or difficulty as a result of filling them out. However, some of them do ask questions about quite personal thoughts and feelings, so we suggest that you complete them somewhere private, where you feel safe and secure, and where you are unlikely to be interrupted. You can take your time filling them out, if this helps, as long as you complete them all with five days of starting them.

If you do find that you are in any way adversely affected by participating in the study, there are a number of sources of help and support available, outlined in the enclosed Debrief Sheet.

Handling of data
In order to be able to analyse the data from the study, we will ask for your consent for members of the research team to have access to your questionnaire responses. All of your completed questionnaire responses will be anonymised by labelling them with a number rather than with your name. They will be stored securely at the IPTS.

What are the possible risks or disadvantages of taking part?
Although it is thought to be unlikely, it is possible that you could experience some emotional distress as a result of filling out the questionnaires. If you experience anything along these lines, we would encourage you to seek help and support from the sources outlined in the Debrief Sheet.

A possible disadvantage of participating in the study is the inconvenience of the time it takes to complete the questionnaires. These have been kept to a minimum and there is some flexibility in how long you can take to complete them, as outlined above.

What are the potential benefits of taking part?

This is not a study that involves treatment or therapy, so you will not benefit along those lines. However, you may experience a sense of satisfaction from having contributed to a piece of research designed to improve understanding of psychological health and ill-health.

In addition, the research team is happy to give you feedback about the meaning of your questionnaire scores, and the study in general, if you would find this helpful.

What if there is a problem?
Any complaint about the way you have been dealt with during the study or any possible harm you might suffer will be addressed. The detailed information on this is given in Part 2.

Confidentiality –who will know that I am taking part in this study?
All the information about your participation in this study will be kept confidential. The details are included in Part 2.

For further information
If you would like any further information about the trial, please do not hesitate to contact Professor Sue Clarke or Helen Bolderston (01202 584120).

This completes Part 1 of the Information Sheet. If the information in Part 1 has interested you and you are considering participation, please continue to read the additional information in Part 2 before making any decision

Part 2

What will happen if I don't want to carry on with the study?
You are free to withdraw from the study at any stage. You do not have to let anyone know that you do not intend to participate.

What if there is a problem?
It is unlikely that this research will cause you any harm. If you have a concern about any aspect of this study, you should phone Sue Clarke (01202 584120). If you remain unhappy, you have the right to complain to the NHS about any aspects of the way you have been approached or treated during the course of this study. In the event that something does go wrong and you are harmed during the research study there are no special compensation arrangements. If you are harmed and this is due to someone's negligence then you may have grounds for a legal action for compensation against Dorset Healthcare Foundation Trust, but you may have to pay your legal costs.

Confidentiality – who will know that I am taking part in this study?
All information relating to you participating in this study will be securely stored, either on a password-protected NHS computer, or locked in an NHS filing cabinet. No completed questionnaires will be labelled using your name or any other identifiable information. Instead, each questionnaire will be labelled with a unique identification number.

If you consent to participate in the study, your GP will be informed. This is a precautionary measure, as in the very unlikely event of you experiencing emotional discomfort following completing the questionnaires, your GP could be a source of support. They will be given no information beyond the fact that you have agreed to participate.

The only people who will have access to any information about you will be the research team. Your contact details, completed questionnaires and consent forms will be accorded the same degree of care as confidential medical records. They will be kept for up to 15 years – the normal period for confidential research data – after which they will be destroyed. Any electronic versions of this information will be stored on password protected NHS hard-drives.

What will happen to the results of this study?
The data from approximately 90 participants will be statistically analysed and submitted to an academic journal for publication (all data will be reported completely anonymously). Publication of the study will enable clinicians and researchers, both nationally and internationally, to learn from the research and to use the new questionnaire to help develop a better understanding of mental health problems and to develop more effective psychotherapy interventions.

Any participants who would like to know the general outcomes of the study, or who would like feedback on their own questionnaire responses are welcome to contact either Sue Clarke or Helen Bolderston on 01202 584120.

Who is organising the study?
The study has been organised by Professor Sue Clarke and her research team. Sue is a Consultant Clinical Psychologist and is the Head of the Intensive Psychological Therapies Service in Poole. The sponsors of this study will pay members of the research team for evaluating your participation in this study.

LREC Approval
This study has been approved by the Southampton and South West Hampshire Research Ethics Committee A. If you have questions about your rights as a participant in this research, or if you feel that you have been placed at risk, you can contact the Research Ethics Committee at: 1st Floor, Regents Park Surgery, Park Street, Shirley, Southampton, SO16 4RJ, Tel. 023 8036 2466

Summary

1. Participation is voluntary. You have the right to choose not to participate, or to stop participating in the study at any point and without consequence.
2. All the information you provide throughout the trial will be completely confidential.
3. This information sheet is for you to keep, as is the enclosed Debrief Sheet
4. For any further information, please contact the IPTS (01202 584120).

Thank you for taking time to read this information sheet.

Participant Debrief Sheet

Cognitive fusion, defined as taking the content of thoughts about a situation to be the literal reality of the situation, is thought to play a role in the development of some psychological difficulties. This study was designed to develop a new questionnaire to measure cognitive fusion, so that researchers can learn more about the process, and also be able to assess the effectiveness of some psychotherapeutic interventions.

In the unlikely event that you are experiencing strong emotions or uncomfortable thoughts as a result of participating in this study, and would like some support, then please contact either members of the research team, Sue Clarke and Helen Bolderston, who are both experienced clinical psychologists. They can both be contacted on 01202 584120 and will be happy to offer appropriate help and support.

If you are in any way distressed as a result of participating in this study, we would also encourage you to make contact with your CMHT care coordinator or IPTS clinician, and/or your family doctor.

You can also contact the Samaritans via their website: www.samaritans.org The Samaritans UK phone number is: 08457 90 90 90.

This study has been approved by the Southampton and South West Hampshire Research Ethics Committee A. If you have questions about your rights as a participant in this research, or if you feel that you have been placed at risk, you can contact the Research Ethics Committee at: 1st Floor, Regents Park Surgery, Park Street, Shirley, Southampton, SO16 4RJ, Tel. 023 8036 2466

Thank you for participating

Appendix F: Public access websites where Study 2 was hosted

http://psych.hanover.edu/research/exponnet.html

http://www.onlinepsychresearch.co.uk

http://www.socialpsychology.org.uk

Appendix G: Study 2 demographic questionnaire and participant information/debrief sheet

Demographics questionnaire

Please indicate your answers in the spaces provided

1. Are you male or female? M F

2. How old are you? _____

3. What is your ethnic origin? White _____ Asian _____

Black _____ Mixed _____

Other ethnic group _____

4. What country do you live in? _____

5. Have you ever sought treatment for a psychological problem (for example depression, anxiety, relationship difficulties, substance misuse, eating disorder)?

Y N

Predictors of personality functioning in adulthood

Information Sheet

Welcome to my internet survey study. I am Helen Bolderston, a PhD student from the University of Southampton, UK. I am requesting your participation in an online study exploring the relationships between various psychological processes, predictors of personality functioning, and adult personality functioning itself.

Taking part in the study

Taking part in this study involves completing a set of questionnaires, which should take no more than 30 minutes. Some of the questions in these questionnaires relate to personal, potentially sensitive issues such as traumatic experiences in childhood and emotional and other difficulties in adulthood. When you complete these questionnaires, it is important to be somewhere where you feel comfortable and safe, and where you are unlikely to be disturbed.
It is entirely up to you to decide whether you would like to participate in this study or not. Your consent to participate in the study will be assumed by completion of the questionnaires.
If you decide to take part in the study you will be asked to complete a brief set of questions about yourself (e.g. your age and ethnic origin) and six other questionnaires of varying lengths.
You are not obliged to complete all the questionnaires, **and in fact you can stop taking part in the study at any stage. For example, having seen the content of some of the questions, you might decide that you do not wish to answer the questions and that you would prefer to stop taking part in the study. That would be fine.** If you do decide to stop early, I recommend that you press the STOP button. This takes you to a final information page. **Included in this final page is an exercise designed to help you manage any uncomfortable thoughts or feelings that might have arisen as a result of you thinking about the subject matter of the study.**

If you are happy to answer all the questions, please remember to click the CONTINUE button at the bottom of each page.

Risk and benefits for participation

The only risk of participation in this study is that some of the questionnaires are of a personal nature and invite you to reflect on potentially sensitive issues. Therefore, it is possible that you might find answering a small number of the questions upsetting. Although completing these questionnaires will be of no direct benefit to you personally, this study will help us understand more fully the various processes leading to and underlying problematic personality functioning in adult life. This in turn may provide information to help guide future psychological treatment development for people with these kinds of psychological difficulties.

Anonymity

Your participation in this study will be anonymous. All data are treated as confidential and will not be linked to any personally identifying information.

Contact

You may contact Helen Bolderston, the principal investigator, if you have any questions or concerns about the study. This can be done by email: heb1w07@soton.ac.uk. This study is supervised by Professor Bob Remington, who can also be contacted by email: R.E.Remington@soton.ac.uk

Legal

This study has been approved by the Ethics Committee, School of Psychology, University of Southampton (contact details below).
Your participation is voluntary and you may withdraw your participation at any time without consequence.

Giving consent

In consenting, I understand that I am not waiving any legal claims, rights, or remedies. I also understand that the data collected as part of this research project will be kept confidential and that published results of this research project will maintain that confidentiality. I finally understand that if I have any questions about my rights as a participant in this research, or feel that I have been placed at risk, I may contact the chair of the Ethics Committee, School of Psychology, University of Southampton, Southampton, S017 1BJ. Phone: +44 (0)23 8059 5578

I confirm that I am **18 years or older**. I have read the information above. By ticking (checking) the box below and clicking CONTINUE, I give my consent to participate in the study described above:

Predictors of personality functioning in adulthood

Debrief Sheet

Research has identified a number of possible predictors of personality problems in adult life. These include some traumatic experiences in childhood and certain temperamental factors. However, many people who experience these predictor factors do not go on to have personality problems in adult life, and little is known about why some people do go on to have such problems and others do not. This study was designed to shed some light on this question.

If you are experiencing uncomfortable thoughts or emotions as a result of taking part in this study, you might wish to complete the following exercise, designed to help you manage such thoughts and emotions. The exercise takes between 5 and 10 minutes to complete. After the exercise there are also some suggestions of places to get support, if you feel that would be helpful.

Exercise

Please read the following guidance all the way through, to get the general idea of the exercise, then work through it again from the beginning, step-by-step, actually trying out each step as instructed before moving onto the next step.

1. Sit in a fairly upright chair, sitting with your back upright, your feet flat on the ground and your hands resting on the arms of the chair or in your lap. Obviously you will need your eyes open to read the instructions, but it is fine if you decide to close your eyes briefly as you practice each step.

2. Take your attention into your feet, noticing the sensations of contact between the soles of your feet and the solid floor. Perhaps you can notice sensations of temperature, texture, or pressure? Continue with this noticing for a minute or so.

3. Similarly, next take your attention into the parts of your body that are in contact with the chair, perhaps your legs or hands, and again notice the physical sensations of this contact between your body and the solid chair. Again, continue with this noticing for a minute or so.

4. Now redirect your attention to the sensations of breathing in your abdomen. Feel each in-breath and each out-breath as they happen. You do not have to change how you are breathing in any way. The aim is just to notice the sensations of breathing, however they are, as you sit here.

If at any point you notice that your attention has wandered from your breathing (as it surely will - human minds were designed to wander!), just acknowledge what it is that has pulled your attention away and then bring your attention back to your breathing. If it is something uncomfortable that pulled your attention away, maybe label it in your mind by saying to yourself something like 'anxiety is here at the moment' or 'uncomfortable memories are around right now'. Then bring your attention back to your breathing, noticing the details of how this in-breath or this out-breath feels. Continue this part of the

exercise for a couple of minutes, coming back to the breath every time your attention wanders.

5. The exercise is not about making any uncomfortable thoughts or emotions disappear. It is designed to help you use the sensations of breathing as a place to hold your attention, somewhere to bring your attention back to whenever you notice that your mind has wandered. In this way, although there may be uncomfortable thoughts and emotions around for a little while, you are less likely to get caught up in them. You can use your breathing as an anchor, holding you steady in the present moment.
6. When you are ready, finish the exercise, but remember that you can always use your breathing as a place to steady your attention, if that is helpful for you.

Following the exercise, if you are still concerned about any issues that have been raised for you through taking part in this study, or if you are experiencing strong emotions or uncomfortable thoughts and would like some support, then you might want to contact your general practitioner/family doctor.

You can also contact the Samaritans via their website: www.samaritans.org There is a link on the Samaritan's website to similar support options for people living outside of the UK. The Samaritans UK phone number is: 08457 90 90 90.

The UK Mind website can also provide UK participants with details of where to seek help for emotional/psychological/mental health problems: Mind PO Box 277 Manchester M60 3XN Tel. 0845 766 0163 Email: info@mind.org.uk Web: www.mind.org.uk

Finally, if you have any questions about your rights as a participant in this research, or you feel that you have been placed at risk, please contact the Chair of the Ethics Committee, School of Psychology, University of Southampton, Southampton, SO17 1BJ, Tel: +44 (0)23 8059 5578

Thank you for participating

Appendix H: List of 48 self-referential adjectives used in Study 3

Main trials: positive words
Respectful
Responsible
Kind
Optimistic
Imaginative
Humorous
Courageous
Considerate
Competent
Friendly
Generous
Creative
Interesting
Understanding
Honest
Intelligent

Main trials: negative words
Unpleasant
Selfish
Egotistical
Irritable
Impolite
Thoughtless
Unfriendly
Hostile
Ungrateful
Intolerant
Deceitful
Unreliable
Insincere
Incompetent
Offensive
Insolent

Recognition trials: Positive words
Brilliant
Positive
Outstanding
Imaginative
Clever
Skilled
Sociable
Cheerful

Recognition trials: Negative words
Moody
Unreasonable
Immature
Irrational
Belligerent
Unappreciative
Boring
Cowardly

Appendix I: Study 3 Brief Demographics Questionnaire

Please complete this questionnaire by ticking the appropriate options.

1. Are you male or female? Male _____ Female _____

2. How old are you? _____ years old

3. What country were you born in? _____

4. What is your ethnic origin? White _____ Asian _____

Black _____ Mixed _____

Other ethnic group _____

THANK YOU FOR COMPLETING THIS QUESTIONNAIRE

Appendix J: Study 3 consent form, participant information sheet and debrief sheet

Consent Form

Study title: Metacognition and cognitive-behavioural responses to visual stimuli
Researcher name: Helen Bolderston
Study reference:
Ethics reference:

Please initial the box(es) if you agree with the statement(s):

I have read and understood the information sheet (Version 1; 20/06/11) and have had the opportunity to ask questions about the study. ☐

I agree to take part in this research project and agree for my data to be used for the purpose of this study ☐

I understand my participation is voluntary and I may withdraw at any time without my legal rights being affected ☐

Data Protection
I understand that information collected about me during my participation in this study will be stored on a password-protected computer and that this information will only be used for the purpose of this study. All files containing any personal data will be made anonymous.

Name of participant (print name)……

Signature of participant……..

Date……

Information Sheet
Metacognition and cognitive-behavioural responses to visual stimuli

I am Helen Bolderston, a PhD student from the University of Southampton. I am requesting your participation in a study exploring influences on performance on a computer-based memory task.

Taking part in the study

Taking part in this study involves completing some questionnaires, which should take no more than 15 minutes. This will be followed by the rest of the study, which involves a computer-based task, and for some participants, a brief psychological exercise. This part of the study will take between 15 and 45 minutes, depending on which aspect of the study you are assigned to.
It is entirely up to you to decide whether you would like to participate in this study or not. You will be asked to provide written consent for participation. If you do consent, you are still not obliged to complete the study.

Risk and benefits for participation

There is no anticipated risk associated with this study.
Although participating will be of no direct benefit to you personally, this study will help us understand more fully processes that may be implicated in the development and maintenance of good mental health.

Anonymity

All data will be stored in an anonymised form, will be treated as confidential and will not be linked to any personally identifying information.
For psychology students from the University of Southampton: Although I ask for your student ID number in order to award participation credits, it will be deleted when the data have been coded.

The study will be administered and data handled by five third year undergraduate project students from the School of Psychology, supervised by Professor Bob Remington and/or by me. They are:
Esther Akinfenwa
Charlotte Deveson
Fay Roberts
Alison Tama
Rebecca Wastell

Contact

You may contact Helen Bolderston, the principal investigator, if you have any questions or concerns about the study. This can be done by email: heb1w07@soton.ac.uk. This study is supervised by Professor Bob Remington: R.E.Remington@soton.ac.uk

Legal

This study has been approved by the Ethics Committee, School of Psychology, University of Southampton.
If you consent, your participation is voluntary and you may withdraw your participation at any time without consequence. If you consent, you will not be waiving any legal claims, rights, or remedies.
If you have any questions about your rights as a participant in this research, or feel that you have been placed at risk, you may contact the chair of the Ethics Committee, School of Psychology, University of Southampton, Southampton, S017 1BJ. Phone: +44 (0)23 8059 5578.

Debriefing Statement

Metacognition and cognitive-behavioural responses to visual stimuli

The aim of this research was to examine the impact of cognitive fusion, a strong attachment to the literal content of human thought that can reduce psychological flexibility. Everyone experiences this to some extent but some people show more fusion than others. We measured your tendency towards fusion using a questionnaire.

Although fusion may be important in determining how flexible we can be, so far there has been little empirical research testing this idea. Current theory suggests that more fusion will tend to be associated with more avoidance of negative self-referential words, but that such words will be easier to recall if seen recently. This study tested that idea by presenting both negative and positive words and measuring how quickly you moved on after seeing them. Later, we also checked how well you recalled the words you had seen.

In some conditions of the study, we carried out a task designed to lower fusion levels before the memory task. Whether or not you took part in this exercise, a similar distraction exercise, or just did the memory task was determined by the design of the study.

Your data will help us understand more about cognitive fusion.

Once again, any reporting of this study will not include your name or any other identifying characteristics. If you would like a copy of the summary of research findings once the project is completed, please email me. If you have any further questions or concerns please contact me Helen Bolderston: heb1w07@soton.ac.uk

If you have been adversely affected in any way by participating in this study, support is available. You could contact me, as I am an experienced clinical psychologist. Alternatively, you could seek support through the following resources:

- University of Southampton: the University Counselling Service, Nightline, on 023 8059 5236 (free from halls on (78)25236) or visit http://nline.susu.org/
- Non-university option: find a counsellor at www.bacp.org

The study could produce different results if you had been fully aware of what you have just read before you took part. For this reason, it is very important to us that you do not disclose the ideas behind the research to your friends or other students, some of whom are likely to take part in the study later on.

Just to remind you, the data from this study may be handled by any of the following 3rd year undergraduate psychology project students:
Esther Akinfenwa
Charlotte Deveson
Fay Roberts
Alison Tama
Rebecca Wastell

If you would like to find out more about cognitive fusion and psychological flexibility, the following articles provide more detailed information:

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

Hayes, S. C., Luoma, J. B., Bond, F. W., Masuda, A., and Lillis, J. (2006). Acceptance and Commitment Therapy: Model, processes and outcomes. *Behaviour Research and Therapy* 44, 1–25.

If you have questions about your rights as a participant in this research, or if you feel that you have been placed at risk, you may contact the Chair of the Ethics Committee, Department of Psychology, University of Southampton, Southampton, SO17 1BJ. Phone: +44 (0)23 8059 5578.

Appendix K: Study 4: Details of protocol development

The content and structure of sessions was based on the successful ACT group protocol developed by SC and HB for the treatment of patients with treatment resistant mental health problems (see Clark et al., 2012, and Clarke et al., in prep, for more details). That original 16-week protocol had in turn been informed by the self-help book "Get Out Of Your Mind And Into Your Life" (Hayes & Smith, 2005), though the content had been substantially modified to be suitable for UK-based mental health patients and for use in a group setting.

Significant changes were made to the Clarke et al. 16-week protocol to render it more suitable for a patient group with high levels of PD psychopathology. These changes were to both the structure of the intervention and the content. In terms of structural changes, the group was extended to 20 sessions, to allow longer for participants to adjust to this new therapeutic approach, to engage with the work, and begin to make changes in their lives. The new protocol included two review sessions, where no new material was introduced. The sessions were used to revisit topics and exercises already covered, as requested by group participants. Also, two hours of 1-to-1 session time with either therapist was made available to each participant, in addition to the group session time. This time could be used as two, 1-hour sessions, or in the form of more frequent, brief sessions. This 1-to-1 time was included to serve the following functions:

1. To aid transition from DBT (which has both group and individual therapy elements).
2. To offer appropriate support and intervention if increased risk of self-harm or increased suicidal ideation was reported, as well as support with difficult emotions, memories and other private experiences.
3. To allow additional clarification and explanation of new, potentially confusing ACT concepts.
4. To provide the opportunity to catch up on missed session content if participants missed sessions and/or were in danger of dropping out of the trial.

In terms of the content of the sessions, changes were made to the original protocol for treatment-resistant mental health problems, by adding some DBT-

oriented content, and by modifying some of the existing ACT content (see Chapter VII, Sections 7.2.4.5 and 7.2.4.6 respectively, for an outline of these changes).

As well as changing the structure and content of the protocol as outlined above, some modifications were made to the original 16-session protocol based on what was learned from the Clarke et al. RCT. For example, the original protocol had relatively little content directly addressing behavioural change, so this aspect of the intervention was enhanced. Also, it was decided to introduce the topic of values, in a relatively brief and unchallenging way, very early in the life of the group, to enhance patient motivation to engage with the intervention.

APPENDIX L: Studies 4 and 5 NHS Research Ethics Committee Approval

RECEIVED
09 FEB 2010

NHS
National Research Ethics Service
SOUTHAMPTON & SOUTH WEST HAMPSHIRE
RESEARCH ETHICS COMMITTEE (A)

1ST Floor, Regents Park Surgery
Park Street, Shirley
Southampton
Hampshire
SO16 4RJ

Tel: 023 8036 2466
023 8036 3462
Fax: 023 8036 4110

Email: scsha.SWHRECA@nhs.net

CM/sta

05 February 2010

Professor Susan E Clarke
Head of Intensive Psychological Therapies Service (IPTS),
Consultant Clinical Psychologist,
Dorset Healthcare NHS Foundation Trust
IPTS
51a Layton Road
Poole
BH12 2BJ

Dear Professor Clarke

Study Title: **Uncontrolled evaluation of a group-based acceptance and commitment therapy intervention for post-dialectical behaviour therapy clients with personality disorder**
REC reference number: **10/H0502/5**
Protocol number: **1**

Thank you for your letter of 28 January 2010, responding to the Committee's request for further information on the above research and submitting revised documentation.

The further information has been considered on behalf of the Committee by the Vice-Chair.

Confirmation of ethical opinion

On behalf of the Committee, I am pleased to confirm a favourable ethical opinion for the above research on the basis described in the application form, protocol and supporting documentation as revised, subject to the conditions specified below.

Ethical review of research sites

The favourable opinion applies to all NHS sites taking part in the study, subject to management permission being obtained from the NHS/HSC R&D office prior to the start of the study (see "Conditions of the favourable opinion" below).

The Committee has not yet been notified of the outcome of any site-specific assessment (SSA) for the non-NHS research site(s) taking part in this study. The favourable opinion does not therefore apply to any non-NHS site at present. I will write to you again as soon as one Research Ethics Committee has notified the outcome of a SSA. In the meantime no study procedures should be initiated at non-NHS sites.

Conditions of the favourable opinion

The favourable opinion is subject to the following conditions being met prior to the start of the study.

Management permission or approval must be obtained from each host organisation prior to the start of the study at the site concerned.

Appendix M: Studies 4 and 5 consent form, participant information sheet, and GP/CMHT information sheet

CONSENT FORM

Title of the project:

Evaluation of Acceptance and Commitment Therapy for post-DBT clients

Name of Researcher: Prof. Sue Clarke

Please initial each box ↓

1.	I confirm that I have read and understood the information sheet dated 14/12/2009 (version 1) for the above study. I have had the opportunity to consider the information, ask questions by phone and have had these answered satisfactorily	
2.	I understand that my participation is voluntary and that I am free to withdraw at any time without my medical care or legal rights being affected	
3.	I understand that data collected throughout this study may be looked at by members of Prof. Sue Clark's research team. I give permission for these individuals to have access to the data collected from this study, to store and to process it.	
4.	I agree that in order for me to participate in the study, members of Prof. Susan Clarke's research team may have access to my contact details, which will be stored securely on an NHS, password-protected computer or locked in an NHS filing cabinet.	
5.	I agree to take part in pre and post research assessment phases (questionnaires and interviews) as well as therapy.	
6.	I consent to the interviews and therapy sessions being audio-taped.	
7.	I agree for my GP and other healthcare professionals involved in my care to be informed about my participation in this research trial.	
8.	I agree that data obtained from this trial can be used, in anonymous form, for publication and for a PhD thesis.	

_____	_____	_____
Paticipant name	Date	Signature

_____	_____	_____
Researcher name	Date	Signature

Participant information sheet
Evaluation of Acceptance and Commitment Therapy for post-DBT clients

You are being invited to take part in a research study. Before you decide whether you would like to take part, you need to understand why the research is being done and what it would involve. Please take time to read the following information carefully. Talk to others about the study if you so wish.
- Part 1 tells you the purpose of this study and what will happen to you if you take part.
- Part 2 gives you more detailed information about the conduct of the study.
If there is anything that is not clear, or if you would like more information, please do not hesitate to contact Helen Bolderston or Sue Clarke on 01202 584120.

Part 1

What is the purpose of this study?
This is a small scale, pilot research trial, which is designed to make some preliminary investigations into the effectiveness of Acceptance and Commitment Therapy (ACT) for people who have stopped engaging in self-harming behaviours following Dialectical Behaviour Therapy (DBT). This study has been designed by senior clinicians and researchers (Prof. Susan Clarke; Prof. Bob Remington; Helen Bolderston).

Why have I been invited?
We are recruiting people who have received at least 12 months of DBT at the Intensive Psychological Therapies Service (IPTS), who have not self-harmed in the last six months, and who are still experiencing some psychological distress or difficulty. Your DBT therapist thought that you meet these criteria and that you might benefit from this therapy and discussed the research trial with you. You indicated that you would be interested in meeting with one of the therapists/researchers (Sue Clarke and Helen Bolderston) for an assessment interview and to hear more about the study in general.
We will be running two ACT therapy groups and aim to recruit a maximum of 10 clients per group. The first group will start in Spring 2010, followed by the second group, which will start in Autumn 2010.

Do I have to take part?
It is up to you whether you take part or not. If you do decide to take part, you will be asked to sign a consent form.
Having signed the consent form, you are still free to withdraw at any time and without giving a reason. A decision to withdraw, or a decision not to take part, will not affect the standard of care you receive, in any way.

What will happen to me if I take part; what will I have to do?

Before therapy
Initially, you will be invited to an interview with one of the therapists/researchers to discuss the study and to assess if the intervention might be suitable for you. Following the interview, if you are offered a place in the study and decide to accept it, you will be invited for a pre-therapy interview at the IPTS, which will take approximately 45 minutes. You will also be asked to complete a set of questionnaires, which will also take approximately 45 minutes.

Therapy
The therapy will be in the form of a 20-week ACT group. There will also be two one-to-one sessions for each client, during the life of the group. ACT is a relatively new therapy that suggests that when people try to avoid difficult thoughts, memories, feelings and so on, they can get into a battle with themselves, and in fact can end up feeling worse. ACT uses exercises such as mindfulness meditation to help people accept difficult thoughts, memories, feelings and so on, so that they can step back from this internal battle. ACT attempts to help people clarify what they really value in life, and to begin to take action in accordance with these values, even if difficult thoughts and feelings are present.

There have been some international ACT studies that have shown promising results in terms of improving the lives of people with difficulties such as depression and anxiety. Over the last few years, there has been research at the IPTS examining the effectiveness of ACT groups for people with entrenched mental health difficulties. This work is on-going, but to date, the majority of patients have experienced a reduction in anxiety and depression levels, as well as an increase in quality of life.

After therapy
Immediately after the therapy group has finished, you will be asked to complete another set of questionnaires and come to a second, shorter interview (that will last approximately 20 minutes). In this interview you will be asked for feedback about your experience of the group, including what was helpful and not so helpful. Six months after the end of the group you will be asked to complete a final set of questionnaires and to attend a final interview.

Data and audio taping
In order to be able to analyse the data from the study, we will ask for your consent for members of the research team to have access to your questionnaire responses. All of your completed questionnaire responses will be anonymised by labelling them with a number rather than with your name. They will be stored securely at the IPTS. We will also ask your permission to audio record the therapy and assessment sessions, for therapist supervision and therapy planning purposes.

Restrictions during and after therapy
If you take part, you can continue taking any medication. We ask that for six months after the end of the ACT group you do not attend any other form of therapy. This is routine practice and is referred to as a consolidation phase. After we collect the six month follow-up data, you can consider any other therapy or intervention available, as appropriate.

Attendance
We ask you to come to all scheduled visits and to complete all questionnaires. If you are going to miss a group session, we ask you to let the clinic know beforehand. If you miss four sessions in a row, you will not be able to continue coming to the group. This is to help you not unintentionally drift out of therapy. It also helps group morale.

What are the alternatives for treatment?
You are free to choose not to participate in this research trial. If you do not want to participate in the trial, you will continue to be able to seek other available treatment, as appropriate.

What are the possible risks or disadvantages of taking part?
As with any therapy, you may sometimes feel emotionally distressed. Your well-being will be monitored by the therapists during every session, and they will endeavour to ensure that no one leaves a session significantly distressed. Clients

who struggle with problems that cannot be addressed adequately in the group will be provided with an individual therapy session. The clinicians are both trained and well experienced in running ACT groups.
As with any therapy, it is possible that a participant might actually deteriorate as a result of participating in the study. It is thought that this is highly unlikely, based on the results of similar ACT studies run at the IPTS. Every attempt will be made to ensure that this does not occur.
A possible disadvantage is the inconvenience of the questionnaires and interviews. These have been kept to a minimum and will be done in a way that is as convenient for you as possible. It is also possible, though unlikely, that you might experience some emotional discomfort as a result of completing some of the questionnaires. Support will be available to you in this event.
We will offer you feedback on the questionnaires and interviews at the end of the study.

What are the potential benefits of taking part?
If you are offered a place in one of the ACT groups and decide to participate, you will be offered 20 weeks of ACT-based psychotherapy. ACT is a relatively new and promising therapy, and there is evidence to suggest that participants might benefit from this intervention. This study provides participants with a therapeutic opportunity that is not currently available in the NHS. On the basis of previous research we anticipate that participants may be more able to build a life that is engaging and satisfying for them by the end of the intervention.
Although group based therapy can seem daunting, it has many benefits. For example, you can develop both from active participation and from observation; you have the opportunity to give and receive immediate feedback; and you have the opportunity for support from people who are experiencing similar difficulties. Many patients find that group-based delivery can actually enhance their experience during therapy. Participation in group therapy does not require you to share personal information.
Whilst we expect these groups to be of benefit to you, we cannot guarantee this. Each participant will receive a £5 gift token at the end of the study as a small gesture to acknowledge the time and effort they have put into being a participant.

What if there is a problem?
Any complaint about the way you have been dealt with during the study or any possible harm you might suffer will be addressed. The detailed information on this is given in Part 2.

Confidentiality –who will know that I am taking part in this study?
All the information about your participation in this study will be kept confidential. The details are included in Part 2.

For further information
If you would like any further information about the trial, please do not hesitate to contact Professor Sue Clarke or Helen Bolderston (01202 584120).

This completes Part 1 of the Information Sheet. If the information in Part 1 has interested you and you are considering participation, please continue to read the additional information in Part 2 before making any decision

Part 2

What if relevant new information becomes available?
Members of Sue Clarke's research team are currently monitoring and will continue to monitor the relevant research literature and websites. If any evidence comes to light that there are any adverse effects of this intervention, your clinician will inform you of these details and ask you whether you would like to continue with the trial. If you decide to continue in the study you will be asked to sign an updated consent form. If you decide not to continue with the trial your continuing care will be arranged. It is also possible that, on receiving new information, the clinician feels that it is in your best interests to withdraw from the study. In the unlikely event that this happens, she will explain the reasons and facilitate the continuation of your care. If the study is stopped for any other reason, you will be told why and your continuing care will be arranged.

What will happen if I don't want to carry on with the study?
You are free to withdraw from treatment at any stage. If you withdraw, we will need to use the data collected up to your point of withdrawal, but this will only be available to members of the research team and will not be stored with information that can identify you. With your permission, we would also like you to complete post-intervention questionnaires and attend the interview despite you not completing the group. However, you will retain the right not to do this if you so choose.

What if there is a problem?
It is unlikely that this therapy will cause you any harm. Trained clinicians will be available at every stage of your involvement. If you have a concern about any aspect of this study, you should phone Sue Clarke (01202 584120). If you remain unhappy, you have the right to complain to the NHS about any aspects of the way you have been approached or treated during the course of this clinical trial. In the event that something does go wrong and you are harmed during the research study there are no special compensation arrangements. If you are harmed and this is due to someone's negligence then you may have grounds for a legal action for compensation against Dorset Healthcare NHS Foundation Trust, but you may have to pay your legal costs.

Confidentiality – who will know that I am taking part in this study?
All information relating to you participating in this study will be securely stored, either on a password-protected NHS computer, or locked in an NHS filing cabinet. No completed questionnaires will be labelled using your name or any other identifiable information. Instead, each questionnaire will be labelled with a unique identification number.
If you consent to participate in the study, your GP and other health professionals involved in your care will be informed. We will ask for your consent to do this.
The only people who will have access to your data from the study will be the research team.

Who is organising the study?
The study has been organised by Professor Sue Clarke and her research team. Sue is a Consultant Clinical Psychologist and is the Head of the Intensive Psychological Therapies Service in Poole.
The sponsors of this study will pay members of the research team for evaluating your participation in this study.

LREC Approval
This study has been approved by the Southampton A Research Ethics Committee

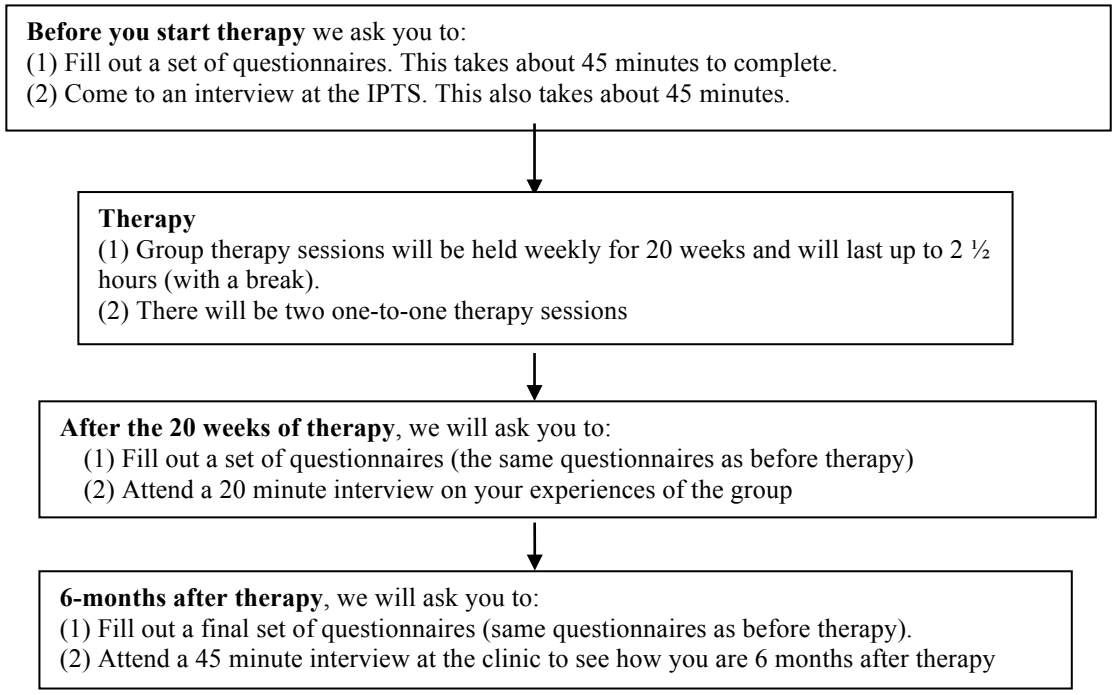
and by the University of Southampton School of Psychology Ethics Committee. If you have questions about your rights as a participant in this research, or if you feel that you have been placed at risk, you can contact the Chair of the Ethics Committee, Department of Psychology, University of Southampton, Southampton, SO17 1BJ. Phone: (023) 8059 3995, or you can contact the Research Ethics Committee at: 1st Floor, Regents Park Surgery, Park Street, Shirley, Southampton, SO16 4RJ, Tel. 023 8036 2466.

Thank you for taking time to read this information pack.

Summary

- **Participation is voluntary. You have the right to choose not to participate, or to stop participating in the trial at any point and without consequence.**
- **All the information you provide throughout the trial will be completely confidential. However, if a member of the team is given reason to believe that you may harm yourself or others, confidentiality may be breached.**
- **This information sheet is for you to keep. If you decide to participate, you will also be provided with a copy of the signed consent form.**
- **For any further information, please contact The IPTS (01202 584120).**

The following flowchart outlines what you will be asked of you if you decide to take part:



GP and CMHT Study Information Sheet

Evaluation of Acceptance and Commitment Therapy for post-DBT clients

Research Background

Personality disorders are common, difficult to treat disorders that can have a profoundly negative impact on the sufferer. Dialectical Behaviour Therapy (DBT; Linehan, 1993) is currently the treatment of choice for many personality disorder clients, enabling many clients to reduce life-threatening behaviours. However, once behaviourally stable, many DBT graduates continue to have a range of psychological difficulties, and live limited and unfulfilling lives.

Acceptance and Commitment Therapy (ACT; Hayes et al., 1999) has shown promise with clients with severe, entrenched mental health problems and may represent an effective post-DBT intervention for behaviourally stable personality disorder clients. For example, Clarke, Kingston et al., (in prep) found that in an RCT comparing ACT with treatment-as-usual for treatment resistant clients (some of whom had personality disorder diagnoses), immediately post therapy participants in both conditions showed improvement on outcome measures of depression, general psychological distress and quality of life. However, at six-month follow-up, the mean score of the control group had deteriorated, whereas in the mean score of the ACT group continued to show improvement.

Research Aim

Despite promising outcomes generally, data from Clarke, Kingston et al. suggests that the hardest-to-help clients remained those previously diagnosed with personality disorders. Although the Clarke, Kingston et al. protocol showed variable outcomes for this group, the results showed enough promise to justify an adaptation of the protocol to target problems associated severe personality disorder.

The purpose of the current study is to develop and test the effectiveness of this new, adapted ACT-based group protocol for clients with personality disorder who have become behaviourally stable through DBT. Although an uncontrolled pre-post intervention, this study is an essential first step to developing an effective and acceptable treatment for this client group, allowing as it does, the opportunity to refine an ACT protocol specifically for personality disorder, and to fine-tune our process and outcome measures and determine effect sizes in preparation for a Randomised Control Trial (RCT).

Methodology

Participants

Participants for this study will be recruited from the Intensive Psychological Therapies Service (IPTS), a specialist NHS personality disorder service based in Poole.

Inclusion criteria

(i) 18 years old and above.

(ii) Personality disorder diagnosis at initial IPTS assessment.

(iii) Minimum of 12 months DBT input prior to this study.

(iv) No parasuicidal behaviour during the six months prior to the study (parasuicidal behaviour defined as deliberate self-harm with or without intent to kill oneself).

(iv) Continued significant psychological difficulties.

Exclusion criteria

(i) Under the age of 18.

(ii) Currently, or in the six months prior to this study, engaged in parasuicidal behaviour.

(iii) Currently meets DSM-IV diagnostic criteria for a psychotic disorder.

(iv) Eating disorder with BMI currently below 17.5.

(v) Currently meets DSM-IV diagnostic criteria for substance dependence.

(iv) Learning disability.

(v) Other organic disorder that might impair capacity to give informed consent, or to participate.

Procedure

Participants will be recruited via DBT therapists at the IPTS. Assessments will be carried out by the CI, Prof. Sue Clarke, or Helen Bolderston, both experienced clinical psychologists. The assessment session will evaluate the client's suitability for the study, and provide the client with the opportunity to discuss the study in detail. If they are offered a place in the study, and if they decide to participate, participants will be asked to provide written consent.

Participants will attend a 20-week ACT group, as well as being offered two one-to-one sessions during the life of the group. Group sessions will be weekly, each will last for 2 ½ hours (including a short break), and will be based at the IPTS. The aims of the group intervention will be to help participants:

(i) Maintain behavioural stability (in terms of not engaging in parasuicidal behaviours)

(ii) Begin to develop a life that they value, which is more satisfying and less defined by mental health problems.

Participants will be asked to complete a questionnaire pack pre-intervention, post-intervention, and at six-month follow-up. They will also have a personality disorder diagnostic interview pre-intervention and at six-month follow-up.

Ethical considerations

As with all forms of psychotherapy, it is possible that some participants might experience emotional discomfort or distress at times during the ACT group intervention. The (very experienced) therapists will monitor the wellbeing of

participants in every group session and will vary the content and intensity of interventions accordingly.

As with all psychotherapy interventions, it is possible that a participant might actually deteriorate as a result of participating in the study. It is thought that this is unlikely, based on previous outcomes with a similar client group in a study by the same research group (Clarke, Kingston et al., in prep). Every attempt will be made to ensure that this does not occur, through weekly monitoring of DBT skills use and risk issues, through directly addressing any participant difficulties as soon as they arise, and through the use of one-to-one sessions where needed. All participants will continue to have access to whatever general forms of support (such as community mental health team care-coordinators) they were receiving prior to the ACT intervention.

It is possible that participants might experience some fatigue, or feel inconvenienced, by completing the questionnaire packs and attending the interviews. The questionnaires included have been kept to the minimum to evaluate the intervention as efficiently as possible. The participant information sheet will include an estimate of the amount of time it will take to complete the questionnaires (approximately 45 minutes), and participants will be informed that it is acceptable to complete the questionnaires over a small number of days, rather than in one sitting. Participants will be offered the opportunity to receive feedback on their questionnaire responses at the end of the study.

Other ethical concerns are limited to issues from research practice such as informed consent, right to withdraw, use of personal data and confidentiality. These will be addressed through research governance practices as follows. Participants will receive full information about all procedures regarding these issues before recruitment and will take part having provided written consent of their participation and data use. It will be emphasised that even following consent, participants are free to withdraw from the study at any point, with no negative consequences.

All data will be labelled with a unique identification number for each participant. Questionnaires will not have any participant identifying information on them and will be kept locked in a filing cabinet at the IPTS. The only people who will have access to any material from participants will be the researchers. All participant completed questionnaires and consent forms will be accorded the same degree of care as confidential medical records.

Following consent from participants, other NHS professionals involved in the care of each individual will be informed about the study. In line with good NHS practice, they will also receive brief clinical reports after assessment, post-intervention, and at discharge at six-month follow-up.

This study is being conducted by Professor Sue Clarke, Consultant Clinical Psychologist, Head of the Intensive Psychological Therapies Service, Dorset Healthcare NHS Foundation Trust. This trial has been approved by the Southampton and South West Hampshire Research Ethics Committee A.

Thank you for taking your time to read this information.

If you have any questions please do not hesitate to contact either Sue Clarke or Helen Bolderston (01202 584120)

Appendix N: Study 4: Details of treatment protocol

The first of the six treatment phases was designed to support participants in the transition from DBT to ACT, and to motivate them to engage with the new ACT content and style. Group members were encouraged to reflect on the ways in which they had benefitted from DBT, and why it was important to continue using DBT skills to maintain stability. The possibility was introduced of a life that is not completely dominated by self-harm urges and behaviour, and each participant was supported to think about a life that had personal meaning for them. ACT was introduced as a possible way of enabling participants clarify what steps in the direction of a personally-valued life might look like, and then gradually and safely to take those first steps. Common ground between ACT and DBT was identified, such as mindfulness as a central process, attention to both acceptance and change, and the basic structure of the weekly group session. At the same time, differences in style (e.g. slower pace and less directive therapists), content (new skills to learn and new ways of thinking about personal experience to consider), and therapeutic goals (more emphasis on a life beyond self-harm) were also identified.

The second phase of treatment addressed creative hopelessness, a common theme early in ACT protocols. In this phase, participants were offered a simple, RFT-consistent account of the costs and benefits of human language and cognitive capacities, particularly in relation to psychological suffering. The universal and wholly understandable coping strategy of EA was discussed, with group members reflecting on the personal, short-term advantages and long-term disadvantages of over-reliance on EA. Willingness was introduced as a possible alternative to EA, with group discussions about how to distinguish when to use each strategy, based on current situation and emotional resources. A ‘dipping your toe in the water of acceptance rather than diving in head first’ approach was encouraged, to help deal with understandable nervousness about ‘having to’ ‘give up’ EA, the predominant psychological coping strategy for all group members.

The third phase of the protocol focused on cognitive defusion. Through several experiential practices as well as didactic teaching, the possibility of cultivating a radically different relationship to cognitive processes was introduced. Group members were encouraged to become more aware of their thoughts as they were arising, something that requires the willingness, at least momentarily, to ‘turn

towards' these private experiences rather than attempting to avoid them. Participants were then coached to view thoughts as just thoughts, regardless of content; to relate to them as mental events that may or may not be accurate reflections of reality. The ACT principle of workability was used to help participants consider how helpful or not it might be to take seriously the content of particular thoughts, particularly thoughts about themselves.

The fourth treatment phase built on this exploration of cognitive defusion by using mindfulness, willingness and self-as-context as ways of cultivating more accepting and effective relationships with private experiences in general (not just thoughts), aversive life history, and sense of self. Experiential exercises and exploration of metaphors such as the chessboard representation of self-as-context were used to support participants to develop a sense of self not merely as the sum of self-referential thoughts. The on-going experience of this sense of self aided the maintenance of a less identified and less avoidant stance towards aversive memories of difficult historic experiences and related emotions.

Having dedicated three months to the gradual cultivation of the ability to relate to private experiences in a less reactive and avoidant manner, the protocol next revisited values, and specifically taking values-consistent action in life. The rationale for this order of content was that although ACT is a behavioural therapy, in that values-consistent behavioural change is the primary aim, a fused and avoidant relationship with thoughts, emotions, sensations, urges, and self can act as a barrier to such behavioural change, no matter how much the individual wishes to see the quality of their life improve. The ability to respond in a defused and open way to the thoughts, emotions and so on that arise as the individual begins to attempt behavioural change, is seen as vital to the individual's ability to persist in taking such action.

A number of written exercises were included to aid group members clarify which aspects of life mattered to them personally, goals and actions related to these valued aspects of life, and the particular barriers preventing them take valued behavioural steps. This work was followed with values-consistent behavioural activation and exposure practices, which in fact formed the substantial part of the homework for the remaining weeks of the group. Emphasis in the protocol continued to be on making 'wise-mind' decisions about whether to use an acceptance or change

approach in each situation, and all behavioural change should be in the form of small, manageable steps.

The final three sessions of the protocol involved reviewing and revising the main themes of the material covered in the intervention, particularly the relationship between DBT and ACT, and acceptance and change, as well as planning for life after the group. The protocol emphasised that straying from valued life directions and falling back into old habits of over-reliance on avoidance and behavioural inertia were deeply human, pretty much guaranteed, and not reasons to give up. Instead, emphasis was placed on the value of ‘starting again’; the capacity to notice when you are no longer living life as you had planned or wanted, to use whatever DBT and ACT strategies are relevant to steady yourself emotionally and address the barriers that have got in the way, and then to begin again planning and carrying out values-consistent steps, no matter how small.

Appendix O: Life Activities Schedule

Please look through this list of activities and indicate which ones you have engaged in over the *last week*, and how many times you have done each activity. Please do this by using the following scale:

1	2	3	4
Once in the last week	More than once in the last week	Every day in the last week	More than once a day in the last week

For example, if you have been to the beach once each day in the last week, you would write 3 by the side of that activity, in the rating column.

If you have not done a particular activity in the last week, please do not write anything by the side of that activity.

Excursions

1. Going on a trip or holiday _____
2. Shopping, car boot sales, flea markets _____
3. Going to the beach _____
4. Going on a picnic _____
5. Going out to dinner _____
6. Going for a drive for pleasure _____
7. Riding in an airplane, hot air balloon, helicopter _____
8. Staying at a hotel or b & b _____
9. Camping _____
10. Going to a museum or exhibit _____
11. Going to a library or bookstore _____
12. Going to a fair, carnival, circus, zoo or amusement park _____
13. Other: _____

Interactions with others/ social activities

1. Going to or giving a party _____
2. Giving or receiving physical affection _____
3. Reminiscing, talking about old times _____
4. Group activities _____
5. Having a frank and open conversation _____
6. Getting together with friends _____
7. Discussing a topic of interest _____
8. Having family visit or visiting family _____
9. Meeting someone new _____
10. Eating out with friends or associate _____
11. Visiting friends or having friends visit _____
12. Other: _____

Entertainment

1. Watching TV or DVDs, or listening to the radio _____
2. Bingo, gambling, playing the lottery (including on-line) _____
3. Going to the cinema _____
4. Going to concerts _____
5. Reading for pleasure _____
6. Going to a play, musical, comedy show _____
7. Going to a sporting event _____
8. Going to the races _____
9. Other: _____

Sports and games

1. Swimming, snorkelling or scuba diving _____
2. Cycling, skating or roller-blading _____
3. Jogging, hiking, or walking _____
4. Going to the gym _____
5. Racquet sports (tennis, badminton, squash, table tennis, racquetball, handball, volleyball) _____
6. Computer or phone games _____
7. Playing board games _____
8. Playing card games _____

9. Puzzles, crosswords, Sudoku etc. _____
10. Golf or putting _____
11. Fishing _____
12. Bird watching _____
13. Rock climbing or mountaineering _____
14. Boating, canoeing, kayaking, sailing _____
15. Pool, snooker or billiards _____
16. Hunting or shooting _____
17. Others: _____

Hobbies, crafts, and the arts

1. Playing a musical instrument _____
2. Singing _____
3. Dancing _____
4. Craft and artwork (drawing, painting, sculpting, pottery, movie making) _____
5. Needlework (knitting, crocheting, embroidery) _____
6. Writing stories, novels, plays, poetry, essays, reports etc _____
7. Restoring antiques or refinishing furniture _____
8. Photography _____
9. Woodworking or carpentry _____
10. Collecting things _____
11. Other: _____

Education

1. Learning something new (a language, how to play a musical instrument etc) _____
2. Learning something artistic (painting, pottery, crocheting etc) _____
3. Educational reading _____
4. Taking a course on something interesting _____
5. Reading a 'how to do it' book or article _____
6. Going to a lecture or to listen to a speaker _____
7. Going back to school, college etc. _____
8. Taking a course in computers _____
9. Other: _____

Health and appearance

1. Getting new clothes, shoes or jewellery _____
2. Putting on makeup or purchasing it _____
3. Getting haircut, going to the hairdressers _____
4. Getting a manicure or pedicure _____
5. Getting a massage _____
6. Putting on perfume or cologne _____
7. Improving one's health (e.g. having teeth fixed, new glasses, eating healthier, starting an exercise programme, having a health check) _____
8. Getting a makeover or facial _____
9. Other: _____

Pampering self and other leisure activities

1. Making free time for yourself _____
2. Playing with or having a pet _____
3. Meditating or doing yoga _____
4. Taking a bubble bath or soothing bath _____
5. Making time to be alone _____
6. Writing a journal or diary or keeping a scrapbook or photo album _____
7. Having a lie in _____
8. Reading the newspaper or magazine _____
9. Listening to music _____
10. Sunbathing _____
11. Enjoying nature _____

12. Telling and listening to jokes _____
13. Going to a spa _____
14. Daydreaming _____
15. Walking barefoot in the sand _____
16. Sitting around a fire _____
17. Other: _____

Domestic activities

1. Cleaning the house _____
2. Baking _____
3. Cooking _____
4. Gardening _____
5. Washing the car, maintaining the car _____
6. Sewing _____
7. Buying flowers and plants _____
8. Re-arranging a room or the house _____
9. Painting and decorating _____
10. Freshening up the house with potpourri, flowers, scented candles etc _____
11. Fixing things around the house _____
12. Other: _____

Treats

1. Chocolate _____
2. Favourite sweets _____
3. Ice cream _____
4. Dessert _____
5. A favourite drink _____
6. Favourite meal _____
7. Other: _____

Altruistic acts

1. Volunteering for a special cause _____
2. Charity work _____
3. Doing favours for others _____
4. Making contributions to religious, charitable or other groups _____
5. Giving gifts _____
6. Helping or listening to someone _____
7. Defending or protecting someone _____
8. Other: _____

Religious activities

1. Going to a place of worship _____
2. Attending a wedding, baptism, bar mitzvah, religious ceremony or function _____
3. Joining a prayer or spiritual group _____
4. Praying _____
5. Reading sacred works _____
6. Participating in a religious fellowship function _____
7. Other: _____

Miscellaneous pleasant activities

1.
2.
3.

Miscellaneous pleasant activities

4.
5.
6.

REFERENCES

Alden, L. (1989). Short-term structured treatment for avoidant personality disorder. *Journal of Consulting and Clinical Psychology, 57,* 756-764.

Altman, D. G. (1996). Better reporting of randomized controlled trials: the CONSORT statement. *British Medical Journal, 313,* 570-1.

Altman, D. G., Schulz, K., Moher, D., Egger, M., Davidoff, F., Elbourne, D., et al., (2001). The revised CONSORT statement for reporting randomized trials: Explanation and elaboration. *Annals of Internal Medicine, 134*, 663-693.

American Psychiatric Association. (2000). *Diagnostic and Statistical Manual of Mental Disorders, 4th edition, text revision.* Washington DC: Author.

Anderson, J. C., & Gerbing, D. W. (1984). The effect of sampling error on convergence, improper solutions, and goodness of fit indices for maximum likelihood confirmatory factor analysis. *Psychometrika*, 49, 155-173.

Anestis, M. D. (2010). What do we know about psychodynamic therapy? A closer look at Shedler's (in press) review. [Web log post]. Retrieved from http://www.psychotherapybrownbag.com/psychotherapy_brown_bag_a/2010/01/what-do-we-know-about-psychodynamic-therapy-a-closer-look-at-shedlers-in-press-review.html

Argyle, M., Trower, P., & Bryant, B. (1974). Explorations in the treatment of personality disorders and neuroses, by social skills training. *British Journal of Medical Psychology, 47,* 211-232.

Arksley, H., & Knight, P. (1999). *Interviewing for social scientists.* London: Sage.

Bach, P., & Hayes, S. C. (2002). The use of Acceptance and Commitment Therapy to prevent the rehospitalization of psychotic patients: A randomized controlled trial. *Journal of Consulting and Clinical Psychology, 70*(5), 1129-1139.

Baer, R. A., Smith, G. T., & Allen, K. B. (2004). Asessment of mindfulness by self-report. The Kentucky Inventory of Mindfulness Skills. *Assessment, 11(3),* 191-206.

Baer, R. A., Smith, G. T., Lykins, E., Button, D., Krietemeyer, J., Sauer, S., et al. (2008). Construct validity of the five facet mindfulness questionnaire in meditating and nonmeditating samples. *Assessment, 15*, 329-342.

Ballinger, C., Yardley, L., & Payne, S. (2004). Observation and action research. In D. F. Marks, & L. Yardley, (Eds), (2004). *Research methods for clinical and health psychology.* London: Sage.

Bandura, A. (1974). Behavior theory and the models of man. *American Psychologist, 29,* 859-869.

Baron, R. M., & Kenny, D. A. (1986). The moderator–mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology, 51,* 1173–1182.

Bastien, J. S., Hermann, B. A., & Moore, D. C. (2010, June). *Mindfulness and acceptance-based group treatment for services professionals with complex trauma: Program description and preliminary evaluation.* Paper presented at 8th Annual World Conference of the Association for Contextual Behavioral Science, Reno, NV.

Bateman, A., & Fonagy, P. (1999). Effectiveness of partial hospitalization in the treatment of borderline personality disorder: A randomized controlled trial. *American Journal of Psychiatry, 156,* 1563-1569.

Bateman, A., & Fonagy, P. (2001). Treatment of borderline personality disorder with psychoanalytically oriented partial hospitalization: An 18-month follow-up. *American Journal of Psychiatry, 158,* 36-42.

Bateman, A., & Fonagy, P. (2009). Randomized Controlled Trial of Outpatient Mentalization-Based Treatment Versus Structured Clinical Management for Borderline Personality Disorder. *American Journal of Psychiatry, 166,* 1355-64.

Beck, A. T. (1976). *Cognitive therapy and the emotional disorders.* New York: International Universities Press.

Beck, A. T., Freeman, A., & Associates. (1990). *Cognitive therapy of personality disorders.* New York: Guildford Press.

Beck, A. T., Steer, R. A., & Brown, G. K. (1996). *Beck Depression Inventory manual (2nd. ed).* San Antonio, TX: Psychological Corporation.

Beckley, K. (2010). Schema therapy. *Contemporary Psychotherapy, 2(2).* Retrieved from http://contemporarypsychotherapy.org/vol-2-no-2/

Bentler, P. (1990). Comparative fit indexes in structural models. *Psychological Bulletin, 107,* 238-246.

Berking, M., Neacsiu, A., Comtois, K. A., and Linehan, M. M. (2009). The impact of experiential avoidance on the reduction of depression in treatment for borderline personality disorder. *Behaviour Research and Therapy*, 47, 663–670.

Bernstein, D. P., Fink, L., Handelsman, L., & Foote, J. (1994). Initial reliability and validity of a new retrospective measure of child abuse and neglect. *American Journal of Psychiatry*, 151 (8), 1132-1136.

Black, D. W., Blum, N., Pfohl, B., & St. John, D. (2004). The STEPPS group treatment program for outpatients with Borderline Personality Disorder. *Journal of Contemporary Psychotherapy, 34*, 193–210.

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

Block, J. A. (2002). *Acceptance or change of private experiences: A comparative analysis in college students with public speaking anxiety.* Doctoral dissertation. University at Albany, State University of New York.

Blum, N., St. John, D., Pfohl, B., Stuart, S., McCormick, B., Allen, J., Arndt, S., Black, D. W. (2008). Systems training for emotional predictability and problem solving (STEPPS) for outpatients with borderline personality disorder: a randomized controlled trial and 1-year follow-up. *American Journal of Psychiatry*, *165*, 468–478

Bohus, M., Haaf, B., Stigmayr, C., Pohl, U., Bohme, R., & Linehan, M. (2000). Evaluation of inpatient dialectical behavioural therapy for borderline personality disorder: A prospective study. *Behavioural Research and Therapy, 38(9),* 875-887.

Bollen, K. A. (1989). *Structural equations with latent variables: Wiley series in probability and mathematical statistics.* New York: Wiley.

Bollen, K. A. and R. A. Stine (1990). Direct and indirect effects: classical and bootstrap estimates of variability. In *Sociological Methodology*, C. Clogg, (Ed), American Sociological Association, Washington, DC.

Bond, F. W., & Bunce, D. (2002). Mediators of change in emotion-focused and problem-focused worksite stress management interventions. *Journal of*

Occupational Health Psychology, 5, 156-163.

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

Boutron, I., Moher, D., Altman, D. G., Schulz, K., & Ravaud, P. (2008). Extending the CONSORT Statement to randomized trials of nonpharmacologic treatment: explanation and elaboration. *Annals of Internal Medicine*, ***148****(4),* 295-309.

Bowen, S.W., Witkiewitz, K., Dillworth, T. M., Chawla, N., Simpson, T., Ostafin, B. D., et al. (2006). Mindfulness mediation and substance use in an incarcerated population. *Psychology of Addictive Behaviours*, 20, 343-347.

Boyatzis, R. E. (1998). *Transforming qualitative information: Thematic analysis and code development.* CA: Sage Publications, Inc.

Bradley, F., Wiles, R., Kinmonth, A., Mant, D., & Gantley, M. (1999). Development and evaluation of complex interventions in health services research: case study of the Southampton heart integrated care project (SHIP). The SHIP Collaborative Group. *British Medical Journal, 318(7185*), 711-5.

Brannen, J. (2005). Mixing Methods: The Entry of Qualitative and Quantitative Approaches into the Research Process. *International Journal of Social Research Methodology, 8*, 173-184.

Brazier, J., Tumur, I., Holmes, M., Ferriter, M., Parry, G., Dent-Brown, K. & Paisley, S. (2006). Psychological therapies including dialectical behaviour therapy for borderline personality disorder: a systematic review and preliminary economic evauation. *Health Technology Assessment, 10* (35).

Bryant, F., Yarnold, P., & Grimm, L. G. (1996). Towards a measurement model of the Affect Intensity Measure: A three-factor structure. *Journal of Research in Personality,* 30, 223-247.

Brown, T. A. (2006). *Confirmatory factor analysis for applied research*. New York: The Guildford Press.

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

Byrne, B. M. (2010). *Structural equation modelling with AMOS.* NY: Routledge.

Campbell, D. T., & Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin,* 56, 81-105.

Campbell, M., Fitzpatrick, R., Haines, A., Kinmonth, A. L., Sandercock, P., Spiegelhalter, P., & Tyrer, P. (2000). Framework for design and evaluation of complex interventions to improve health. *British Medical Journal, (321),* 694-696.

Cattel, R. B. (1966). *Handbook of multivariate experimental psychology.* Chicago: Rand McNally.

Chambless, D. M., & Hollon, S. (1998). Defining empirically supported therapies. *Journal of Consulting and Clinical Psychology, 66*, 7-18.

Chanen, A. M., Jackson, H. J., McCutcheon, L.K., Dudgeon, P., Jovev, M., Yuen, H. P., et al. (2008). Early intervention for adolescents with borderline personality disorder: a quasi-experimental comparison with treatment as usual. *Australian and New Zealand Journal of Psychiatry*, 43 (5), 397 – 408.

Christensen, L., & Mendoza, J.L. (1986). A method of assessing change in a single subject: An alteration of the RC index. *Behavior Therapy, 17*, 305-308.

Ciarrochi, J., Bilich, L., & Godsell, C. (2010). Psychological flexibility as a mechanism of change in acceptance and commitment therapy. In R. A. Baer (Ed.), *Assessing mindfulness and acceptance processes in clients.* (pp. 51-76). Context Press.

Clark, D. A., & Beck, A. T., with Alford, B. A. (1999). *Scientific Foundations of Cognitive Theory and Therapy of Depression.* Philadelphia: Wiley & Sons

Clark, L. A., & Watson, D. (1995). Constructing validity: Basic issues in objective scale development. *Psychological Assessment, 7,* 309-319.

Clark-Carter, D., & Marks, D. F. (2004). Intervention studies: design and analysis. In D. F. Marks, & L. Yardley, (Eds.), *Research methods for clinical and health psychology.* London: Sage.

Clarke, G. N. (2003). Improving the Transition From Basic Efficacy Research to Effectiveness Studies: Methodological Issues and Procedures. In A. E. Kazdin (Ed.), *Methodological Issues and Strategies in Clinical Research.* APA.

Clarke, S. E., Kingston, J., & Remington, R., Wilson, K. G., and Bolderston, H. (2012). Acceptance and Commitment Therapy (ACT) for a Heterogeneous

Group of 'Treatment Resistant' Clients: A Treatment Development Study. *Cognitive and Behavioral Practice,* 19, 560-572.

Clarke, S. E., Kingston, J., & Remington, R., Wilson, K. G., and Bolderston, H. (2013). An RCT comparing ACT and treatment as usual for treatment-resistant mental health patients. Manuscript in preparation.

Clarke, S. E., Thomas, P., & James, K. (2012). A Randomised Trial of the Efficacy of Cognitive Analytic Therapy (CAT) for the Treatment of Personality Disorder. *British Journal of Psychiatry,* Advance online publication. doi: 10.1192/bjp.bp.112.108670

Clarkin, J. F., Levy, K. N., Lenzenweger, M. F., & Kernberg, O. F. (2004). The Personality Disorders Institute/Borderline Personality Disorder Research Foundation randomized control trial for borderline personality disorder: Rationale, methods, and patient characteristics. *Journal of Personality Disorders, 18,* 52–72.

Clarkin, J. F., Levy, K. N., Lenzenweger, M. F., & Kernberg, O. F. (2007). Evaluating three treatments for borderline personality disorder: A multiwave study. *American Journal of Psychiatry,* 164*,* 922–928.

Cochrane, A., Barnes-Holmes, D., Barnes-Holmes, Y., Stewart, I. & Luciano, C. (2007). Experiential avoidance and aversive visual images – response delays and event-related potentials on a simple matching task. *Behaviour Research and Therapy,* 45, 1379-1388.

Cohen, J. (1962). The statistical power of abnormal-social psychological research: a review. *Journal of Abnormal and Social Psychology, 65,* 145-53.

Cohen J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). New Jersey: Erlbaum.

Coid, J., Yang, M., Tyrer, P., Roberts, A., & Ullrich, S. (2006). Prevalence and correlates of personality disorder in Great Britain. *British Journal of Psychiatry,* 188, 423-431.

Comtois, K. A., Elwood, L., Holdcraft, L. C., Smith, W. R., & Simpson, T. L. (2007). Effectiveness of dialectical behavior therapy in a community mental health center. *Cognitive and Behavioral Practice, 14,* 406–414.

Cortina, J. M. (1993). What is coefficient alpha? An examination of theory and applications. *Journal of Applied Psychology, 78,* 98-104.

Cramer, V., Torgersen, S. & Kringlen, E. (2007). Socio-demographic conditions,

subjective somatic health, Axis I disorders and personality disorders in the common population: The relationship to quality of life. *Journal of Personality Disorders*, **21**, 552–56

Cronbach, L.J., & Meehl, P.E. (1955). Construct validity in psychological tests. *Psychological Bulletin, 52*, 281-302.

Crowell, S. E., Beauchaine, T. P., & Linehan, M. (2009). A biosocial developmental model of borderline personality: Elaborating and extending Linehan's theory. *Psychological Bulletin, 135*, 495-510.

Davidson, K. M. (2000). *Cognitive therapy for personality disorders: A guide for clinicians.* 2nd edition. London: Arnold (Hodder).

Davidson, K. M., Gumley, A., Millar, H., Drummond, L., & Macaulay, F. (2006a). A randomized controlled trial of cognitive behavior therapy for borderline personality disorder: Rationale for trial, method, and description of sample. *Journal of Personality Disorder, 20,* 431-449.

Davidson, K. M., Norrie, J., Tyrer, P., Gumley, A., Tata, P., Murray, H., et al. (2006b). The effectiveness of Cognitive Behavior Therapy for borderline personality disorder: Results from the borderline personality disorder study of Cognitive Therapy (BOSCOT) trial. *Journal of Personality Disorder, 20,* 450-465.

Davidson, R. J., Kabat-Zinn, J., Schumacher, J., Rosenkranz, M., Muller, D., Sanotelli, S. F., et al. (2003). Alterations in brain and immune function produced by mindfulness meditation. *Psychosomatic Medicine, 65,* 564-570.

Denman, C. (2001). Cognitive-analytic therapy. *Advances in Psychiatric Treatment, 7,* 243-256.

Derogatis, L. R. (1994). *Symptom Checklist-90-R (SCL-90-R): Administration, scoring and procedures manual.* Minneapolis, MN: National Computer Systems Inc.

DeRubeis, R. J., Evans, M.D., Hollon, S. D., Garvey, M. J., Grove, W. M., & Tuason, V. B. (1990). How does cognitive therapy work? Cognitive change and symptom change in cognitive therapy and pharmacotherapy for depression. *Journal of Consulting and Clinical Psychology, 58*, 862-869.

Diener, E., Wirtz, D., Tov, W., Kim-Prieto, C., Choi, D., Oishi, S., et al., (2009). New measures of well-being: Flourishing and positive and negative feelings. *Social Indicators Research,* 39, 247-266.

Dimeff, L., & Linehan, M. M. (2001). Dialectical behavior therapy in a nutshell. *The California Psychologist,* 34, 10-13.

Dimidjian, S., Hollon, S. D., Donson, K. S., Schmaling, K. B., Kohlenberg, R. J., Addis, M. E., Gallop, R., McGlinchey, J. B., Markley, D. K., Gollan, J. K., Atkins, D. C., Dunner, D. L., & Jacobson, N. S. (2006). Randomized trial of behavioural activation, cognitive therapy, and antidepressant medication in the acute treatment of adults with major depression. *Journal of Consulting and Clinical Psychology,* 74, 658-670.

Dobkin, P. L. (2008). Mindfulness-based stress reduction: What processes are at work? *Complementary Therapies in Clinical Practice, 14,* 8-16.

Doering, S., Horz, S., Rentrop, M., Fischer-Kern, M., Schuster, P., Benecke, C., et al. (2010). Transference-focused psychotherapy v. treatment by community psychotherapists for borderline personality disorder: randomised controlled trial. *British Journal of Psychiatry, 196*(5), 389 - 395.

Dolan-Sewell, R. T., Krueger, R. F., & Shea, M. T. (2001). Co-occurrence with syndrome disorders. In W. J. Livesley (Ed), *Handbook of personality disorders* (pp. 84–104). New York: Guilford Press.

Duggan, C., Huband, N., Smailagic, N., Ferriter, M. and Adams, C. (2007), The use of psychological treatments for people with personality disorder: A systematic review of randomized controlled trials. Personality, 1: 95–125.

Edwards, P. (2010). Questionnaires in clinical trials: guidelines for optimal design and administration, *Trials*, 11, p. 2-10.

Eid, M., & Deiner, E. (2006). Introduction: The need for multimethod measurement in psychology. In M. Eid & E Deiner (Eds), *Handbook of multimethod measurement in psychology* (pp. 3-8). Washington DC: American Psychological Association.

Fabrigar, L. R., Wegener, D. T., McCallum, R. C., & Strahan, E. J. (1999). Evaluating the use of exploratory factor analysis in psychological research. *Psychological Methods, 4,* 272-299.

Farrell, J. M., Shaw, I., & Webber, M. A. (2009). A schema-focused approach to group psychotherapy for outpatients with borderline personality disorder: a randomized controlled trial. *Journal of Behavior Therapy and Experimental Psychiatry, 40,* 317–328.

Ferguson, E. (2005). The use of diary methodologies in health and clinical psychology. In J. Miles & P. Gilbert (eds.) *A handbook of research methods in clinical and health psychology* (pp. 175-184). Oxford: Oxford University Press.

Fernandez-Ballesteros, R., & Botella, J. (2008). Self-report measures. In A. M. Nezu, & C. M. Nezu, (Eds), *Evidence-based outcome research: A practical guide to conducting randomized controlled trials for psychosocial interventions.* (pp. 95- 122). Oxford University Press.

Field, A. P. (2005). *Discovering Statistics Using SPSS.* London: Sage.

Field, A. P., & Davey, G. C. L. (2005). Experimental Methods in Clinical Research. In J. Miles & P. Gilbert (eds.) *A handbook of research methods in clinical and health psychology* (pp. 175-184). Oxford: Oxford University Press.

Field, A. P., & Hole, G. (2003). *How to design and report experiments*. London: Sage.

Flett, G. L., & Hewitt, P. L. (1995). Criterion validity and psychometric properties of the Affect Intensity Measure in a psychiatric sample. *Personality and Individual Differences,* 19, 585-591.

Forman, E. M., Herbert, J. D., Moitra, E., Yeomans, P. D., Geller, P. A. (2007). A randomized controlled effectiveness trial of acceptance and commitment therapy and cognitive therapy for anxiety and depression, *Behavior Modification,* 31, 772-799.

Forrester, M. A. (Ed) (2010). *Doing Qualitative research in Psychology: A Practical Guide*. London: Sage.

Frith, H., & Gleeson, K. (2004). Clothing and embodiment: men managing body image and appearance. *Psychology of men and masculinity*, 5, 40-48.

Fylan, F. (2005). Semi-structured interviewing. In J. Miles & P. Gilbert (eds.), *A handbook of research methods in clinical and health psychology* (pp. 175-184). Oxford: Oxford University Press.

Garber, M. A., Nau, D. P., Erickson, S. R., Aikens, J.E., & Lawrence, J. B. (2004). The Concordance of Self-Report with Other Measures of Medication Adherence: A Summary of the Literature. *Medical Care, 42*, 649-652.

Gaudiano, B. A. (2009). Öst's (2008) Methodological Comparison of Clinical Trials of Acceptance and Commitment Therapy versus Cognitive Behavior Therapy: Matching Apples with Oranges? Behaviour Research and Therapy, 47, 1066-1070.

Gaudiano, B. A., & Herbert, J. D. (2006). Acute treatment of inpatients with psychotic symptoms using Acceptance and Commitment Therapy: Pilot results. *Behaviour Research & Therapy, 44*(3), 415-437.

Gaudiano, B. A., Miller, I. W., & Herbert, J. D. (2007). The treatment of psychotic major depression: Is there a role for adjunctive psychotherapy? *Psychotherapy and Psychosomatics, 76*, 271-277.

Giesen-Bloo, J. G., van Dyck, R., Spinhoven, P., van Tilburg, W., Dirksen, C., van Asselt, T., et al. (2006). Outpatient psychotherapy for borderline personality disorder: Randomized trial of schema-focused therapy vs. transference-focused psychotherapy. *Archives of General Psychiatry, 63,* 649– 658.

Gilbert, P., & Irons, C. (2005). Thinking about research: issues and dilemmas. In J. Miles & P. Gilbert (eds.) *A handbook of research methods in clinical and health psychology* (pp. 175-184). Oxford: Oxford University Press.

Gillanders, D. T., Bolderston, H., Bond, F. W., Dempster, M., Flaxman, P. E., Campbell, L., et al., (2013). The development and initial validation of the Cognitive Fusion Questionnaire. Manuscript submitted for publication.

Goldstein, T. R., Axelson, D. A., Birmaher, B., & Brent, D. A. (2007). Dialectical behavior therapy for adolescents with bipolar disorder: A 1-year open trial. *Journal of American Academic Child and Adolescent Psychiatry, 46,7,* 820-830.

Grant, B. F., Hasin, D. S., Stinson, F. S., Dawson, D. A., Chou, S. P., Ruan,W. J., et al. (2004). Prevalence, correlates, and disability of personality disorders in the United States: Results from the national epidemiological survey on alcohol and related conditions. *Journal of Clinical Psychiatry, 65*(7), 948–958.

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. (2006) Risk factors for deliberate self-harm among female college students: The role and interaction of childhood maltreatment, emotional inexpressivity, and affect intensity/reactivity. *American Journal of Orthopsychiatry,* 76(2), 238–250.

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., & Roemer, L. (2008). The relationship between emotion dysregulation and deliberate self-harm among female undergraduate students at an urban commuter university. *Cognitive Behaviour Therapy, 37*, 14-25.

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.

Greco, L. A., Lambert, W., & Baer, R. A. (2008). Psychological inflexibility in childhood and adolescence: development and evaluation of the Avoidance and Fusion Questionnaire for Youth, *Psychological Assessment,* 20(2), 93-102.

Gregory RJ; Chlebowski S; Kang D; Remen AL; Soderberg MG; Stepkovitch J et al. (2008). A controlled trial of psychodynamic psychotherapy for co-occurring borderline personality disorder and alcohol use disorder. *Psychotherapy, 45*(1): 28-41.

Gunderson, J. G., & Gabbard, G. O. (1999). Making the case for psychoanalytic therapy in the current psychiatric environment. *Journal of the American Psychoanalytic Association, 47,* 679–704.

Harley R, Sprich S, Safren S, Jacobo M, Fava M (2008). Adaptation of dialectical behavior therapy skills training group for treatment-resistant depression. *Journal of Nervous and Mental Disease*, Feb; 196(2):136-43.

Harned, M. S., Chapman, A. L., Dexter-Mazza, E. T., Murray, A., Comtois, K. A., & Linehan, M. M. (2008). Treating co-occurring Axis I disorders in recurrently suicidal women with borderline personality disorder: A 2-year randomized trial of dialectical behavior therapy versus community treatment by experts. *Journal of Consulting and Clinical Psychology, 76,* 1068–1075.

Harrington, D. (2009). *Confirmatory Factor Analysis.* New York: Oxford University Press, Inc.

Hayes, S. C. (2004). Acceptance and Commitment Therapy and the new behavior therapies: Mindfulness, acceptance and relationship. In S. C. Hayes, V. M. Follette, and M. Linehan (Eds.), *Mindfulness and acceptance: Expanding the cognitive behavioral tradition* (pp. 1-29). New York: Guilford.

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

Hayes, S. C., Luoma, J. B., Bond, F. W., Masuda, A., and Lillis, J. (2006). Acceptance and Commitment Therapy: Model, processes and outcomes. *Behaviour Research and Therapy* 44, 1–25.

Hayes, S. C., & Strosahl, K. D. (Eds.) (2004). *A practical guide to Acceptance and Commitment Therapy.* New York: Springer.

Hayes, S. C., Strosahl, K. D., & Wilson, K. G. (1999). *Acceptance and Commitment Therapy: An Experiential approach to behavior change*. New York: Guildford Press.

Hayes, S. C., Strosahl, K. D., Wilson, K. G., Bissett, R. T., Pistorello, J., Toarmino, D., et al. (2004). Measuring experiential avoidance: A preliminary test of a working model. *The Psychological Record, 54,* 553-578.

Hayes, S C., Wilson, K. G., Gifford, E. V., Bissett, R., Piasecki, M., Batten, S. V., et al. (2004). A preliminary trial of Twelve-Step Facilitation and Acceptance and Commitment Therapy with polysubstance-abusing methadone-maintained opiate addicts. *Behavior Therapy, 35, 667-688.*

Haynes, S. N. (2003). Clinical Applications of Analogue Behavioral Observation: Dimensions of Psychometric Evaluation. In A. E. Kazdin (Ed.), *Methodological Issues and Strategies in Clinical Research.* APA.

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

Herman, J. L., Perry, J. C., & Van der Kolk, B. A. (1989). Childhood trauma in borderline personality disorder. *American Journal of Psychiatry,* 146, 490-495.

Hollis, S., & Campbell, F. M. (1999). What is meant by 'intention to treat' analysis? *British Medical Journal*, *319*, 670-4.

Hollon, S., & Kendall, P. C. (1980). Cognitive self-statements in depression: Development of an automatic thoughts questionnaire. *Cognitive Therapy and Research, 4,* 383-395.

Howitt, D., & Cramer, D. (2005). *Introduction to Statistics in Psychology.* London: Pearson/Prentice Hall.

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling, 6,* 1–55.

Hurley, M. A., & Holmes, E. P. (2010, July). Contextually based DBT: Description and initial findings. Poster session presented at ACBS World Conference VIII, Reno, NA.

Hunt, G. M., and Azrin, N. H. (1973). Community reinforcement approach to alcoholism. *Behaviour Research & Therapy,* 11, 91-104.

Hyler, S. E., Rieder, R. O., Williams, J. B. W., Spitzer, R. L., Lyons, M., & Hendler, J. (1988). The Personality Diagnostic Questionnaire: Development and preliminary results. *Journal of Personality Disorders, 2,* 229-237.

Iacobucci, D., (2010). Structural equation modeling: Fit indices, sample size, and advanced issues. *Journal of Consumer Psychology*, 20, 90-98.

Jacobson, N. S., Dobson, K. S., Truax, P. A., Addis, M. E., Koerner, K. Gollan, J., et al. (1996). A component analysis of cognitive-behavioral treatment for depression. *Journal of Consulting and Clinical Psychology, 64,* 295-304.

Jacobson, N. S., & Truax, P. (1991). Clinical significance: A statistical approach to defining meaningful change in psychotherapy research. *Journal of Consulting and Clinical Psychology, 59,* 12-19.

Johnson, J. G., Cohen, P., Brown, J., Smailes, E. M., and Bernstein, D. (1999). Childhood Maltreatment Increases Risk for Personality Disorders During Early Adulthood. *Archives of General Psychiatry,* 56(7), 600-606.

Johnson, J. G., Cohen, P., Chen, H., Kasen, S., and Brook, J. S. (2006). Parenting Behaviors Associated With Risk for Offspring Personality Disorder During Adulthood. *Archives of General Psychiatry,* 63(5), 579–587.

Judd, C. M., & Kenny, D. A. (2010). Data analysis. In S. Fiske, D. Gilbert, G. Lindsay (Eds.), *The handbook of social psychology*. Boston: McGraw-Hill.

Kaiser, H. F. (1970). A second-generation little jiffy. *Psychometrika,* 35, 401-415.

Kaiser, H. F. (1974). An index of factorial simplicity. *Psychometrika, 39, 31-36.*

Kashdan, T.B., Barrios, V., Forsyth, J. P., & Steger, M. F. (2006). Experiential avoidance as a generalized psychological vulnerability: Comparisons with coping and emotion regulation strategies. *Behaviour Research and Therapy,* 44, 1301-1320.

Kashdan, T.B. & Steger, M.F. (2006). Expanding the topography of social anxiety: An experience sampling assessment of positive emotions and events, and emotion suppression. *Psychological Science, 17,* 120-128.

Kazdin, A. E. (2003). *Methodological Issues and Strategies in Clinical Research.* APA.

Kazdin, A. E. (2007). Mediators and mechanisms of change in psychotherapy research. *Annual Review of Clinical Psychology, 3*, 1-27.

Kendler, K. S., Aggen, S. H., Czajkowski, N., Røysamb, E., Tambs, E., Torgersen, S., Neale, M. C., and Reichborn-Kjennerud, T. (2008). The Structure of Genetic and Environmental Risk Factors for *DSM-IV* Personality Disorders: A Multivariate Twin Study. *Archives of General Psychiatry,* 65(12), 1438-1446.

Kincaid, J.P., Fishburne, R.P., Rogers, R.L., & Chissom, B.S. (1975). Derivation of New Readability Formulas (Automated Readability Index, Fog Count, and Flesch Reading Ease formula) for Navy Enlisted Personnel. Research Branch Report 8-75. Chief of Naval Technical Training: Naval Air Station Memphis.

Kingston, J. L., Clarke, S., & Remington, R. (2010). Experiential avoidance and problem behaviour: A meditational Analysis. *Behavior Modification,*34(2), 145-163.

Kline, P. (1986). *A handbook of test construction.* London: Methuen.

Kline, R. B. (2005). *Principles and practice of structural equation modeling* (2nd ed.). New York: The Guildford Press.

Koons, C.R., Robins, C.J., Tweed, J.L., Lynch, T.R., Gonzalez, A.M., G.K., Morse, J.Q., Bishop, G.K., Butterfield, M.I., & Bastian, L.A. (2001). Efficacy of

Dialectical Behavior Therapy in Women Veterans with Borderline Personality Disorder. Behavior Therapy, 32, 371 – 390.

Kraemer, H. C., Wilson, T., Fairburn, C. G., & Agras, W. S. (2002). Mediators and moderators of treatment effects in randomized clinical trials. *Archives of General Psychiatry, 59,* 877−883.

Krietman, N. (1977). *Parasuicide*. Chichester: Wiley.

Krinsley, K. E., Young, L. S., Weathers, E W., Brief, D. J., & Kelley, J. M. (1992). Behavioral correlates of childhood trauma in substance abusing men. Paper presented at the annual meeting of the Association for Advancement of Behavior Therapy, Boston, MA.

Kuyken, W., Byford, S., Taylor, R. S., Watkins, E., Holden, E., White, K., et al., (2008). Mindfulness-Based Cognitive Therapy to Prevent Relapse in Recurrent Depression. *Journal of Consulting and Clinical Psychology, 76,* 966-978.

Lambert, M. J. (1994). Use of psychological tests for outcome assessment. In M. E. Maruish, (Ed.), *The use of psychological testing for treatment planning and outcome assessment* (pp. 75-97). Hillsdale, NJ: Erlbaum.

Lamont, S., and Brunero, S. (2009). Personality Disorder Prevalence and Treatment Outcomes: A Literature Review. *Issues in Mental Health Nursing*, 30, 631–637.

Lancaster, G., Dodds, S., & Williamson, P. (2004). Design and analysis of pilot studies: Recommendations for good practice. *Journal of Evaluation in Clinical Practice, 10,* 307-312.

Larsen, D. L., Attkisson, C. C., Hargreaves, W. A., & Nguyen, T. D. (1979). Assessment of client/patient satisfaction: Development of a general scale. *Evaluation and Program Planning*, *2*, 197-207.

Larsen, R.J., & Diener, E. (1987). Affect intensity as an individual difference characteristic: A review. *Journal of Research in Personality*, *21*, 1-39.

Leichsenring, F. (2005). Are psychodynamic and psychoanalytic therapies effective? *International Journal of Psychoanalysis, 86,* 841-868.

Leichsenring, F. & Leibing, E. (2003). The effectiveness of psychodynamic therapy and cognitive behavior therapy in the treatment of personality disorders: a meta analysis. *American Journal of Psychiatry*, 160,1, 223–32.

Leichsenring, F. & Leibing, E. (2007). Supportive-Expressive (SE) Psychotherapy: An Update. *Current Psychiatry Reviews, 3,* 57-64.

Lejuez, C. W., Hopko, D. R., & Hopko, S. D. (2001) A Brief Behavioral Activation Treatment for Depression : Treatment Manual. Behavior Modification, 25, 255-286.

Lenzenweger, M. F., & Clarkin, J. F. (Eds). (2005). *Major Theories of Personality Disorder.* New York, London: The Guildford Press. 2nd Ed.

Levin, M., & Hayes, S. C. (2009). Is Acceptance and Commitment Therapy Superior to Established Treatment Comparisons? *Psychotherapy and Psychosomatics*, 78, 380-381.

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

Linehan, M. M., Armstrong, H. E., Suarez, A., Allmon, D., & Heard, H. L. (1991). Cognitive-behavioral treatment of chronically parasuicidal borderline patients. *Archives of General Psychiatry, 48,* 1060-1064.

Linehan, M.M., Comtois, K.A., Murray, A.M., Brown, M.Z., Gallop, R.J., Heard, H. L. et al. (2006). Two-year randomized trial + follow-up of Dialectical Behavior Therapy vs. therapy by experts for suicidal behaviors and borderline personality disorder, Archives of General Psychiatry, 63, 757 – 766.

Linehan, M. M., Dimeff, L. A., Reynolds, S. K., Comtois, K. A., Welch, S. S., Heagerty, P., & Kivlahan, D. R. (2002). Dialectical behavior therapy versus comprehensive validation therapy plus 12-step for the treatment of opioid dependent women meeting criteria for borderline personality disorder. *Drug and Alcohol Dependence, 67(1),* 13-26.

Linehan, M. M., Heard, H. L., & Armstrong, H. E. (2003). Naturalistic follow-up of a behavioral treatment for chronically parasuicidal borderline patients. Archive of General Psychiatry, 50(12), 971-974.

Linehan, M. M., Schmidt, H., III, Dimeff, L. A., Craft, J. C., Kanter, J., & Comtois, K. A. (1999). Dialectical behavior therapy for patients with borderline personality disorder and drug-dependence. *American Journal on Addictions, 8,* 279–292.

Linehan, M. M., Tutel, D. A., Heard, H. L., & Armstrong, H. E. (1994). Interpersonal outcome of cognitive behaviour treatment for chronically suicidal borderline patients. *American Journal of Psychiatry,* 151, 1771-1776.

Livesley, W. J., Jang, K. L., and Vernon, P., A. (1998). Phenotypic and genetic structure of traits delineating personality disorder. *Archives of General Psychiatry*, 55(10), 941-948.

Lovaas, O. I., Koegel, R., Simmons, J., & Long, J. (1973). Some generalization and follow-up measures on autistic children in behavior therapy. *Journal of Applied Behaviour Analysis, 6,* 131-165.

Luborsky, L. (1984). Principles of psychoanalytic psychotherapy: A manual for supportive-expressive treatment. Basic Books.

Lynch, T. R. (2000). Treatment of elderly depression with personality disorder comorbidity using dialectical behavior therapy. *Cognitive and Behavioral Practice*, 7, (4), 468-477.

Lynch, T. R., Chapman, A. L., Rosenthal, M. Z, Kuo, J. R., & Linehan, M. M. (2006). Mechanisms of change in dialectical behavior therapy: Theoretical and empirical observations. *Journal of Clinical Psychology, 62*, 459-480.

Lynch, T. R. & Cheavens, J. S. (2008), Dialectical behavior therapy for comorbid personality disorders. J. Clin. Psychol., 64: 154–167

Lynch, T. R., Cheavens, J. S., Cukrowicz, K. C., Thorp, S. R., Bronner, L., Beyer, J. (2006). Treatment of older adults with co-morbid personality disorder and depression: A dialectical behavior therapy approach. *International Journal of Geriatric Psychiatry,* 22(2), 131–143.

Lynch, T. R., Cheavens, J. S., Morse, J. Q., & Rosenthal, M. Z. (2004). A model predicting suicidal ideation and hopelessness in depressed older adults: the impact of emotion inhibition and affect intensity. *Aging and Mental Health*, 8(6), 486-497.

Lynch, T. R., Morse, J. Q., Mendelson, T., & Robins, C. J. (2003). Dialectical behavior therapy for depressed older adults: a randomized pilot study. *Am. J. Geriatric Psychiatry*, 11:33–45.

Lynch, T. R., Robins, C. J., Morse, J. Q., & Krause, E. D. (2001). A mediational model relating affect intensity, emotion inhibition, and psychological distress. *Behavior Therapy,* 32, 519-536.

Lynch, T. R., Trost, W. T., Salsman, N., Linehan, M. M. (2007). Dialectical behavior therapy for borderline personality disorder. *Annual Review of Clinical Psychology*, 3, 181–205.

Lynch, T. R., et al. (2011 – 2016). Dialectical behavior therapy for treatment-resistant depression: A randomised controlled trial. Efficacy and mechanism evaluation programme. MRC, NIHR.

MacKinnon, D.P., & Fairchild, A.J. (2009). Current directions in mediation analysis. *Current Directions in Psychology*, 18, 16-20.

MacKinnon, D. P., Lockwood, C. M., Hoffman, J. M., West, S. G., & Sheets, V. (2002). A comparison of methods to test mediation and other intervening variable effects. *Psychological Methods*, *7*, 83-104.

MacKinnon, D. P., Warsi, G., & Dwyer, J. H. (1995). A simulation study of mediated effect measures. *Multivariate Behavioral Research*, *30*, 41-62.

Manias, E., Botti, M., & Bucknall, T. (2002). Observation of pain assessment and management – the complexities of clinical practice. *Journal of Clinical Nursing, 11*, 724-733.

Mann, C. J. (2003). Observational research methods. Research design II: cohort, cross sectional and case-control studies. *Emergency Medical Journal, 20,* 54-60.

Marks, D. F. (2004). Questionnaires and surveys. In D. F. Marks, & L. Yardley, (Eds), (2004). *Research methods for clinical and health psychology.* London: Sage.

Marks, D. F., & Yardley, L. (Eds), (2004). *Research methods for clinical and health psychology.* London: Sage.

Marsh, H. W. (1996). Positive and negative global self-esteem: A substantively meaningful distinction or artifactors? Journal of Personality and Social Psychology, 70, 810-219.

Marx, B. P., & Sloan, D. M. (2002). The Role of Emotion in the Psychological Functioning of Adult Survivors of Childhood Sexual Abuse. *Behavior Therapy*, 33, 563-577.

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

Masuda, A., Twohig, M. P., Stormoa, A. R., Feinstein, A. B., Chou, Y., & Wendell, J. W. (2010). The effects of cognitive defusion and thought distraction on emotional discomfort and believability of negative self-

referential thoughts. *Journal of Behavior Therapy and Experimental Psychiatry, 41,* 11.17.

McCracken, L. M., Vowles, K. E., & Eccleston, C. (2005). Acceptance-based treatment for persons with complex, long-standing chronic pain: A preliminary analysis of treatment outcome in comparison to a waiting phase. *Behaviour Research and Therapy,* 43, 1335–1346.

McGlashan, T. H., Grilo, C. M., Skodol, A. E., Gunderson, J. G., Shea, M. T., Morey, L. C., Zanarini, M. C., Stout, R. L. (2000). The Collaborative Longitudinal Personality Disorders Study: baseline Axis I/II and II/II diagnostic co-occurrence. *Acta Psychiatrica Scandanavia,* 102, 256-264.

McMain, S. F., Links, P. S., Gnam, W. H., Guimond, T., Cardish, R. J., Kormann, L., & Streiner, D. L. (2009). A randomized trial of dialectical behavior therapy versus general psychiatric management for borderline personality disorder. *American Journal of Psychiatry, 166,* 1365–1374.

Millon, T. Millon, C. & Davis, R. (1994) 'Millon Clinical Multiaxial Inventory-III Manual'. Minneapolis, MN: National Computer Systems Inc.

Mogg, K., Bradley, B. P., Field, M., & De Houwer, J. (2003). Eye movements to smoking-related pictures in smokers: relationship between attentional biases and implicit and explicit measures of stimulus valence. *Addiction,* 98, 825-836.

Monroe-Blum, H., & Marziali, E. (1995). A controlled trial of short-term group treatment for borderline personality disorder. *Journal of Personality Disorders,* 9, 190-198.

Muran, J. C., Safran, J. D., Samstag, L. W., & Winston, A. (2005). Evaluating an alliance-focused treatment for personality disorders. *Psychotherapy: Theory, Research, Practice, Training, 42*(4), 532-545.

National Institute for Health and Clinical Excellence (2009). *Borderline Personality Disorder – The NICE guideline on treatment and management*. NICE

National Institute for Health and Clinical Excellence (2010). *Antisocial Personality Disorder – The NICE guideline on treatment and management*. NICE

National Institute for Mental Health in England. (2003). *Personality disorder: No longer a diagnosis of exclusion. London*: National Institute for Mental Health.

Neff, K. D. (2003). The development and validation of a scale to measure self-compassion. *Self and Identity,* 2, 223-250.

Nezu, A. M., & Nezu, C. M. (Eds), (2008). *Evidence-based outcome research: A practical guide to conducting randomized controlled trials for psychosocial interventions.* Oxford University Press.

Nock, M. K., Janis, I. B., & Wedig, M. M. (2008). Research designs. In Nezu, A. M., & Nezu, C. M. (Eds). *Evidence-based outcome research: A practical guide to conducting randomized controlled trials for psychosocial interventions.* Oxford University Press.

Nunnally, J. C. (1970). *Introduction to psychological measurement*. New York: McGraw-Hill.

Nunnally, J. C. (1978). *Psychometric theory* (2nd ed.). New York: McGraw-Hill.

Nunally, J. C., & Bernstein, I. H. (1994). *Psychometric Theory* (3rd ed.). New York: McGraw-Hill.

Øst, L. G. (2008). Efficacy of the third wave of behavioral therapies: A systematic review and meta-analysis. *Behaviour Research and Therapy, 46,* 296-321.

Paulhus, D. L. (1991). Measurement and control of response bias. In J. P. Robinson, P. R. Shaver & L. S. Wrightsman (Eds.), *Measures of personality and social psychological attitudes* (pp. 17-59). San Diego, CA: Academic Press.

Perry, J. C. (1992). Problems and considerations in the valid assessment of personality disorders. *American Journal of Psychiatry, 149,* 1645-53.

Persons, J. B. & Silberschatz, G. (2003). Are Results of Randomized Controlled Trials Useful to Psychotherapists? In *Methodological Issues and Strategies in Clinical Research.* APA.

Piper, W. E., & Joyce, A. S. (2001). Psychosocial treatment outcome. In W. J. Livesley (Ed), *Handbook of personality disorders.* NY: Guildford.

Plomin, R., and Caspi, A. (1999). Behavioral genetics and personality. In L. Pervin and O. P. John (Eds.), *Handbook of Personality: Theory and research* (2nd ed., pp. 251-276). New York: Guildford Press.

Preacher, K. J., & Hayes, A. F. (2004). SPSS and SAS procedures for estimating the indirect effects in simple mediation models. *Behaviour Research Methods, Instruments and Computers, 36,* 717-737.

Raczek, S. W. (1992). Childhood abuse and personality disorders. *Journal of Personality Disorders, 6,* 109-116.

Reddy, M. K., Pickett, S. M., & Orcutt, H. K. (2006). Experiential avoidance as a mediator in the relationship between childhood psychological abuse and current mental health symptoms in college students. *Journal of Emotional Abuse, 6*, 67-85.

Robins, C. J., & Chapman, A. L. (2004). Dialectical behavior therapy: Current status, recent developments, and future directions. *Journal of Personality Disorders*, 18, 73−79.

Robins, L. N., Helzer, J. E., Croughan, J. L., & Ratcliff, K. S. (1981). National Institute of Mental Health Diagnostic Interview Schedule: its history, characteristics and validity. *Archives of General Psychiatry, 38*, 381-389.

Roger, D. & Najarian, B. (1989). The construction and validation of a new scale for measuring emotion control. *Personality and Individual Differences, 10(8),* 845-853.

Rosenthal, M. Z., Cheavens, J. S., Lejuez, C. W., & Lynch, T. R. (2005). Thought suppression mediates the relationship between negative affect and borderline personality disorder symptoms. *Behaviour Research and Therapy, 43*, 1173-1185.

Rosenthal, M. Z., Polusny, M. A., & Follette, V. M. (2006). Avoidance mediates the relationship between perceived criticism in the family of origin and psychological distress in adulthood. *Journal of Emotional Abuse, 6,* 87-102.

Rounsaville, B.J., Carroll, K.M., & Onken, L.S. (2001). A stage model of behavioral therapies research: Getting started and moving on from Stage I. Clinical Psychology: *Science & Practice, 8*(2), 133-142.

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

Ruiz-Canela M, Martinez-Gonzalez MA, de Irala-Estevez J. (2000). Intention to treat analysis is related to methodological quality. *British Medical Journal, 320*, 1007-8.

Rusch, N., Schiel, S., Corrigan, P. W., Leihener, F., Jacob, G. A., Olschewski, M. et al., (2008). Predictors of dropout from inpatient dialectical behavior therapy

among women with borderline personality disorder. *Journal of Behavior Therapy, 39, 497-503.*

Russell, D. W. (2002). In search of underlying dimensions: The use (and abuse) of factor analysis in Personality and Social Psychology Bulletin. *Personality and Social Psychology Bulletin, 28,* 1629-1646.

Rutter, M. (2004). Developmental catch-up and deficit, following adoption after severe global early privation. *Journal of Child Psychology and Psychiatry,* 39, 465-476.

Ryle, A. (1985). Cognitive theory, object relations and the self. *British Journal of Medical Psychology, 58, 1–*7.

Ryle, A. (1990). *Cognitive Analytic Therapy: Active Participation in Change*. Chichester: John Wiley & Sons.

Ryle, A. (1997). *Cognitive Analytic Therapy and Borderline Disorders: The Model & the Method.* Chichester: J. Wiley & Sons.

Scheel, K.R. (2000). The empirical basis of Dialectical Behavior Therapy: summary, critique and implications. *Clinical Psychology: Science & Practice, 7*(1), 68-86.

Scher, C. D., Stein, M. B., Asmundson, G. J. G., McCreary, D. R., & Forde, D. R. (2001). The childhood trauma questionnaire in a community sample: Psychometric properties and normative data. *Journal of Traumatic Stress*, 14, 843–857.

Segal, Z. V., Williams, J. M. G., & Teasdale, J. T. (2002). *Mindfulness-based cognitive therapy for depression: A new approach to preventing relapse.* New York: Guilford Press.

Seligman, M. E. P. (1995). The effectiveness of psychotherapy. The *Consumer Reports* study. *American Psychologist, 50,* 956–974.

Shafran, R., Thordarson, D., & Rachman, S. (1996). Thought- action fusion in obsessive compulsive disorder. Journal of Anxiety Disorders, 10(5), 379–391.

Shedler, J. (2010). The efficacy of psychodynamic psychotherapy. *American Psychologist, 65,* 98-109.

Skinner, B. F. (1957). Verbal Behavior. New York: Appleton-Century-Crofts.

Skinner, B. F. (1969). *Contingencies of Reinforcement.* New York: Appleton-Century-Crofts.

Skodol, A. E., and Bender, D. S. (2009). The Future of Personality Disorders in DSM-V? *American Journal of Psychiatry*, 166, 388-391.

Sobel, M. E. (1982). Asymptotic intervals for indirect effects in structural equations models. In S. Leinhart (Ed.), *Sociological methodology 1982* (pp. 290–312). San Francisco: Jossey-Bass.

Sperry, L. (1995). Paradigm shift in the treatment of personality disorders. In L. Sperry, *Handbook of Diagnosis and Treatment of the DSM-IV Personality Disorders (*pp. 1-14). New York: Brunner/Mazel.

Spitzer, R. Gibbon, M. & Williams, J. (1996) 'Structured Clinical Interview of DSM IV Axis I and II disorders'. New-York: N.Y. State psychiatric Institute, Biometrics Research Department.

Spokas, M., Rodebaugh, T., & Heinberg, R. (2008). Therapy Outcome. In D. McKay (Ed.), *Handbook of research methods in abnormal and clinical psychology.* California: Sage Publications.

Stravynski, A., & Shahar, A. (1983). The treatment of social dysfunction in nonpsychotic outpatients. A review. *Journal of Nervous and Mental Disorder, 171*(12), 721-8.

Strosahl, K. D., Hayes, S. C., Wilson, K. G., & Gifford, E. V. (2004). An ACT Primer. In Hayes, S. C., & Strosahl, K. D. (Eds.) (2004). *A practical guide to Acceptance and Commitment Therapy.* p 32. New York: Springer.

Svartberg, M., Stiles, T., & Seltzer, M. (2004). Randomized, controlled trial of the effectiveness of short-term dynamic psychotherapy and cognitive therapy for Cluster C personality disorders. *American Journal of Psychiatry, 161*, 810-817.

Sydenstricker-Neto, J. (1997). Research Design and Mixed-Method Approach. A Hands-on Experience. Retrieved from http://www.socialresearchmethods.net/tutorial/Sydenstricker/bolsa.html#Why%20Mixed

Tabachnick, B. J., & Fidell, L. S. (2001). *Using multivariate statistics*. Needham Heights: Allyn & Bacon.

Taylor, G. (2010a). *Understanding and Ameliorating Stigma towards Clients with a Personality Disorder: An Acceptance and Commitment Therapy-based Approach.* Unpublished manuscript, School of Psychology, University of Southampton.

Taylor, G. (2010b). *The Validation of the Believability of Stigmatising Thoughts Questionnaire.* Unpublished manuscript, School of Psychology, University of Southampton.

Teasdale, J. D., Segal, Z. J., Ridgeway, V. A., & Soulby, J. M. (2000). Prevention of relapse/recurrence in major depression by Mindfulness-Based Cognitive Therapy. *Journal of Consulting and Clinical Psychology, 68,* 615-623.

Telch, C.F., Agras, W.S. & Linehan, M.M. (2000). Group Dialectical Behavior Therapy for binge eating disorder: A preliminary uncontrolled trial. Behavior Therapv, 31, 569-582.

Thurstone, L. L. (1947). *Multiple factor analysis.* Chicago: University of Chicago Press.

Torgersen, S., Kringlen, E., & Cramer, V. (2001). The prevalence of personality disorders in a community sample. *Archive of General Psychiatry,* 58(6), 590-596.

Trochim, W. M. K. (2006). *Convergent & Discriminant Validity.* Retrieved from http://www.socialresearchmethods.net/kb/convdisc.php

Turner, R. M. (2000). Naturalistic evaluation of dialectical behavior therapy–oriented treatment for borderline personality disorder. *Cognitive and Behavioral Practice, 7,* 413–419.

Van den Bosch, L. M. C., Verheul, R., Schippers, G. M., & van den Brink, W. (2002). Dialectical behavior therapy of borderline patients with and without substance use problems. Implementation and long-term effects. *Addictive Behaviors, 27,* 911–923.

Verheul, R, Andrea, H., Berghout, C., Dolan, C. C., van Busschbach, J.J., van der Kroft, P.J.A., Bateman, A.W. & Fonagy, P. (2008). Severity Indices of Personality Problems (SIPP-118): Development, factor structure, reliability, and validity. Psychological Assesment, 20, 23- 34.

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 Behavior Therapy for women with borderline personality disorder. British Journal of Psychiatry, 182, 135 – 140.

Verheul, R., & Widiger, T. A. (2004). A meta-analysis of the prevalence and usage of the personality disorder not otherwise specified (PDNOS) diagnosis. *Journal of Personality Disorders,* 18, 309–319.

Vinnars, B., Barber, J.P., Norén, K., Thormählen, B., Gallop, R., & Weinryb, R.M. (2005). Supportive-Expressive Psychotherapy in Personality Disorders: An Outpatient Randomized Controlled Trial. *American Journal of Psychiatry, 16,2* 1933-1940.

West, S.G., Finch, J. F., and Curran, P.J. (1995). Structural equation models with non-normal variables: Problems and remedies. In R. H. Hoyle (Ed.), *Structural equation modelling: Concepts, issues, and applications* (pp. 56-75). Thousand Oaks, CA: Sage.

Westbrook, D., & Kirk, J. (2005). The clinical effectiveness of Cognitive Behavior Therapy: Outcome for a large sample of adults treated in routine practice. *Behavior Research and Therapy, 43,* 1243-1261.

Westen, D., (2000). The efficacy of dialectical behavior therapy for borderline personality disorder. *Clinical Psychology: Science and Practice. 7,* 92-94.

Westen, D., Shedler, J., & Bradley, R. (2006). A prototype approach to personality disorder diagnosis. American Journal of Psychiatry, 163, 846–856.

Westen, D., Novotny, C.M., & Thompson-Brenner, H. (2004). The empirical status of empirically supported psychotherapies: Assumptions, findings, and reporting in controlled clinical trials. *Psychological Bulletin*, 130, 631 – 663.

Wetherall, M., &Vedhara, K. (2005) The measurement of physiological outcomes in health and clinical psychology. In J. Miles & P. Gilbert (eds.) *A handbook of research methods in clinical and health psychology* (pp. 175-184). Oxford: Oxford University Press.

Wicksell, R. K., Reofalt, J., Olsson, G. L., Bond, F., & Melin, L. (2008). Avoidance and cognitive fusion – Central components in pain related disability? Development and preliminary validation of the Psychological Inflexibility in Pain Scale (PIPS). *European Journal of Pain,* 12, 491-500.

Widaman, K. F. (1993). Common factor analysis versus principal component analysis: Differential bias in representing model parameters? *Multivariate Behavioral Research*, *28*, 263-311.

Widiger, T. A. (2008). Clinical interviews. In A. M. Nezu, & C. M. Nezu, (Eds), *Evidence-based outcome research: A practical guide to conducting randomized controlled trials for psychosocial interventions.* (pp. 47- 66). Oxford University Press.

Widiger, T.A., & Frances, A. (1989). Epidemiology, diagnosis, and comorbidity of borderline personality disorder. In A. Tasman, R. Hales, & A. Frances (Eds.), Review of Psychiatry (Vol. 8, pp. 8-24). Washington, DC: American Psychiatric Press.

Widom, C.S., DuMont, K., & Czaja, S.J. (2007). A prospective investigation of major depressive disorder and comorbidity in abused and neglected children grown up. *Archives of General Psychiatry,* 64, 49–56.

Wilkinson, S. (2005). Using focus groups: exploring the meanings of health and illness. In J. Miles & P. Gilbert (eds.) *A handbook of research methods in clinical and health psychology* (pp. 175-184). Oxford: Oxford University Press.

Wilkinson, S., Joffe, H. & Yardley, L. (2004). Qualitative data collection: interviews and focus groups. In D. Marks and L. Yardley (Eds.), *Research methods in clinical and health psychology* (pp. 39-55). London: Sage.

Wilson, D.B., Lipsey, M. W. (2001). The role of method in treatment effectiveness research: Evidence from meta-analysis. *Psychological Methods, 6,* 413-429.

Wilson, K. G., Bond, F., Ciarrochi, J., Hayes, S., & Bordieri, M., (2010). Reconstructing Measurement Theory From a Contextualist Perspective. Panel discussion at the Association for Contextual Behavioral Science, World Congress VIII, Reno, NV.

Wilson, K. G., & DuFrene, T. (2008). *Mindfulness for two: An acceptance and commitment therapy approach to mindfulness in psychotherapy.* Oakland, CA: New Harbinger Publications.

Wilson, K. G., Sandoz, E. K., Kitchens, J., & Roberts, M. E. (2010). The Valued Living Questionnaire: Defining and measuring valued action within a behavioral framework. *The Psychological Record, 60,* 249-272.

Winston, A., Laikin, M., Pollack, J., Samstag, L.W., McCullough, L., & Muran, J.C. (1994). Short-term psychotherapy of personality disorders. *American Journal of Psychiatry*, 151, 190–194.

Wolpe, J. (1958). *Psychotherapy by reciprocal inhibition.* Stanford, CA: Stanford University Press.

Young, J. E. (1990). *Cognitive therapy for personality disorders: A schema focused approach.* Sarasota, FI: Professional Resource Exchange.

Young, J. E. (2004). *Schema Therapy for Borderline Personality Disorder.* Retrieved from http://schematherapy.com/

Zanarini, M. C., Frankenburg, F. R., Hennen, J., Reich, D. B., & Silk, K. R. (2004). Axis I comorbidity in patients with borderline personality disorder: 6-year follow-up and prediction of time to remission. *American Journal of Psychiatry, 161*(11), 2108-14.

Zettle, R. D. & Hayes, S. C. (1986). Dysfunctional control by client verbal behavior: The context of reason giving. *The Analysis of Verbal Behavior, 4,* 30 38.

Zigmond, A.S. & Snaith, R.P. (1983) The Hospital Anxiety and Depression Scale. Acta Psychiatrica Scandinavica, 67, 361-370.

Zimmerman, M., & Coryell, W. H. (1990). *DSM–III* PD dimensions. *Journal of Nervous and Mental Disease, 178,* 686–692.