The Role of Alliance and Symptomatic Change Within Cognitive Behaviour Therapy for Depression

A thesis presented in partial fulfillment of the requirements for the degree of Doctor of Clinical Psychology at Massey University, Albany, New Zealand.

Carol Ann Osborne

2010

ABSTRACT

This thesis explores the role of alliance processes and symptomatic change within Cognitive Behaviour Therapy for depression (A. T., Beck, Rush, Shaw, & Emery, 1979). Archived session data from The Depression Outcome Study conducted at Massey University, Albany (2006-2009) and a single-case research design with multiple assessments was used to determine temporal relations between alliance and depression severity. An observer version of the Working Alliance Inventory-Short Revised (WAI-SR-O) was utilized to rate the alliance of ten client-therapist dyads every session over the first ten sessions of therapy. Symptomatic change was assessed every session with the Beck Depression Inventory–II.

Increasing inter-rater reliability of the current research involved seeking guidelines for rating the WAI from researchers overseas. These guidelines were modified and expanded to rate the WAI-SR-O within CBT. A rater reliability study was conducted in two stages to provide a forum to train the raters for the current research and establish inter-rater reliability.

Similar to previous research, results of the current research demonstrated that an early strong alliance may predict a positive outcome and poor early alliance may lead to premature termination of therapy. It was difficult to draw definite conclusions as to whether alliance precedes symptomatic change. However, findings suggested that a reciprocal relationship between alliance and symptomatic change may start in the assessment stages of therapy.

The current research demonstrated a clear reciprocal relationship between Total Alliance scores and depression severity in some sessions in some cases. A strong alliance contributed to a decrease in depression severity which subsequently increased the alliance. However, in other sessions the reciprocal relationship was not as clear. Furthermore, definite conclusions could not be drawn about the reciprocal effects between the components of the alliance (i.e., Goal, Task and Bond subscales) and depression severity. However, symptomatic change was found to be greater in the context of a strong bond between the client and therapist. There was also some evidence of subscale scores increasing following a decrease in depression severity and decreasing following an increase in depression severity in the same session that the depression severity was rated. Furthermore, there was evidence that findings were related to variability of data, nature of the alliance and time and environmental factors.

ACKNOWLEDGEMENTS

My educational experience has been enriched by many people, all of whom have given their time, support and intellects to encourage and foster my development. Dr Nikolas Kazantzis has been a central figure in this journey. As my supervisor he inspired me and provided many opportunities to professionally develop within this research. I am grateful for his advice, support and enthusiasm, and critical and supportive feedback on my thesis drafts.

Another debt of gratitude is owed to Dr Paul Merrick, my second supervisor, who became my first supervisor when Dr. Kazantzis left Massey University to take up employment at La Trobe University, Melbourne. His breath of knowledge encouraged and inspired me. Thank you Paul for all your ideas, your wisdom and your constant support of my professional endeavors.

A huge thanks to Dr Janet Leathem who objectively reviewed my thesis with "fresh eyes". Janet has been my mentor through most of my post-graduate endeavors and I have greatly appreciated her ongoing strength and support.

This thesis could not have been completed without the ongoing support from family, friends and work colleagues. I thank my mother, Shirley and my two daughters, Kristy and Joanna for patiently enduring this process with me and offering support and encouragement along the way.

Several individuals have also provided professional assistance and essential materials and deserve special recognition. Dr. Adam Horvath graciously permitted the use of the observer version of the Working Alliance Inventory–Short Revised. Dr. Patrick Raue generously sent me guidelines for the Working Alliance Inventory that he and his colleagues developed for their research. He then provided me with much appreciated feedback for the modification and expansion of the guidelines for my own research. I also appreciated the feedback Dr. Lisa Fenton and Dr. Lois Gelfand provided during the development of those guidelines. I would like to acknowledge the professional assistance of Dr. Barry MacDonald and Dr Neville Blampied who supplied me with statistical and analysis knowledge when I needed it. Thank

you also to the extramural library staff who were prompt to send articles and books and source overseas theses when requested.

I was exceedingly fortunate to have Robyn Vertongen as my clinical supervisor in the Depression Outcome Study. I am grateful to have had the opportunity to learn from her.

Thank you to the eight participants of the rater reliability study which further assisted my understanding of the alliance measure within my research and the newly developed guidelines to rate that measure. To Jenni and Lynn, who went on to become the two raters for my research, a special thanks. I appreciated the numerous hours you both spent training and watching and rating DVDs.

I also express my appreciation to fellow post graduate students; Rachel Findlay, Jeanne Daniels, Margo Munro, Michael Easden, Nicole Foster, Robyn Gedye, Cintamani Mallison and Anna Connolly, who were part of an ongoing supervision group in the first years of my research project. I thank you all for your shared ideas, enthusiasm and assistance.

Finally, and most importantly, I thank the participants of the Depression Outcome Study for their willingness to participate in the study and who openly shared their very private experiences with their therapists, thus making this investigation possible.

Without the support of you all this research could not have happened.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Chapter 1: Orientating Framework and Overview of Thesis

Overview

In the ongoing search for ways to improve treatment outcomes, the study of psychotherapy processes has become an important area of research in psychology. It is no longer enough to know that there has been an improvement in symptomology. Now, it is crucial to know *how* individuals change as well (Laurenceau, Hayes, & Feldman, 2007). The therapeutic alliance is one important part of psychotherapy process (Freud, 1910; Gaston, 1990; Rogers, 1951, 1957; Wampold, 2001) and will be the focus of the research described here. It is hoped that advancing the understanding of the role of the alliance process and symptomatic changes in depressive within Cognitive Behaviour Therapy (CBT; A. T. Beck, et al., 1979) for depression will contribute to enhancing the effectiveness this therapy. By studying these critical dimensions of therapy the current research will further inform training and supervision of CBT therapists to work effectively with depressed clients. This chapter presents an orientating framework and rationale for the current research including: brief review of depression, CBT, the study of psychotherapy process, the alliance process and its measurement and temporal design. It also outlines the structure of this thesis.

Depression

The World Health Organisation predicts that by 2020 depression will be second only to cardiovascular disease in terms of worldwide burden of ill health (Murray & Lopaz, 1996). Te Rau Hinengaro: The New Zealand Mental Health Survey (2006) reported the lifetime prevalence (the population of people known to have met the criteria at some time in their lives) for major depression disorder was 16.0% (p < .0001) (Oakley Browne, Wells, & Scott, 2006). For diagnosis of major depression, the Diagnostic and Statistical Manual of Mental Disorders–IV–TR (DSM-IV-TR; American Psychiatric Association, 2000) requires five or more of the following symptoms: depressed mood, pervasive loss of interest or pleasure, significant weight change or change in appetite, sleep disturbance, observable agitation or retardation, loss of energy, feelings of worthlessness or unnecessary guilt, poor concentration or decision making and recurrent thoughts of death or suicide. According to the DSM-IV-TR at least one of the first two symptoms must be present. This pattern must be present most of the day, nearly every day during a continuous period of at least two weeks, and must cause significant distress or impairment in social, occupational or other

important areas of functioning (e.g., personal hygiene) (American Psychiatric Association, 2000; Sadock & Sadock, 2003).

Prevalence of Depression

Major depression can vary in severity and usually lasts six to nine months. When depression lasts longer, and the depressive symptoms go away for a short time, only to return again, it is called chronic depression (Wells, Burnam, Rogers, Hays, & Camp, 1992). At least 20% of people with an initial episode of major depression do not recover within 2 years (Scott, 2001). Recovery means to have several months without meeting the criteria of a major depressive episode (American Psychiatric Association, 2000). Furthermore, long term studies following the recurrence or meeting the criteria of a major depressive episode again (Kennedy, Abbott, & Paykel, 2003; Ramana et al., 1995; Van Londen, Molenaar, Goekoop, Zwinderman, & Rooijmans, 1998) have demonstrated high relapse rates. Raman et al. (1995) found 40% of their subjects relapsed within the first 10 months of remission. Remission means that the person has several weeks with minimal depressive symptoms and no major depressive episode (American Psychiatric Association, 2000). Van Londen et al. (1998) found a recurrence rate of 41% within five years. Individuals with residual symptoms tended to relapse in the first four months after remission, while individuals without residual symptoms recurred mainly after 12 months after remission.

Impact of Depression

Major depression disorder can have a major impact on people's health and their quality of life (Khan-Bourne & Brown, 2003). Furthermore, the costs (e.g., care, work productivity and psychosocial costs) of an individual having this disorder fall upon the individual, their family and friends, employers and to society as a whole (Klerman & Weissman, 1992). It has high comorbidity with other disorders including: panic disorder, agoraphobia, social phobia, generalized anxiety disorder, post traumatic stress disorder and substance abuse, which can make it more difficult to treat (Hirschfeld, 2001; Sadock & Sadock, 2003; Schoevers, Beekman, Deeg, Jonker, & van Tilburg, 2003; Souery et al., 2007). Numerous studies support CBT (A. T. Beck et al., 1979) which emphasizes cognitive and behaviour models, as a viable treatment for individuals with major depressive disorder (Craighead, Sheets, Brosse, & Ilardi, 2007).

Cognitive-Behaviour Therapy

CBT is a brief psychological treatment focused on changing problematic beliefs and behaviours. Cognitive change is hypothesized to lead symptom change (A. T. Beck et al., 1979; J. S. Beck, 1995). CBT is recognized internationally (Lambert & Ogles, 2004) and is popular among psychological practitioners and consumers in New Zealand (Kazantzis & Deane, 1998; Koong Hean Foo & Merrick, 2004). Evolving through the 1960s and 1970s (Dobson & Dozois, 2001), it is also recognized as a well established empirically supported treatment for a broad range of clinical and medical disorders (A. T. Beck, 1997; Butler, Chapman, Forman, & A. T. Beck, 2006; Hollon & A. T. Beck, 1979, 2004; Reinecke & Freeman, 2003; Salkovskis, 1996) including: major depression disorder (Butler et al., 2006; Fava et al., 2004; Hollon et al., 2005; Hollan & A. T. Beck, 2004, Scott, Palmer, Paykel, Teasdale, & Hayhurst, 2003), bipolar disorder (Lam et al., 2000; Lam et al., 2003; Scott, Garland, & Moorehead, 2001), eating disorders (Agras et al., 1992; Leitenberg et al., 1994; Wilson & Fairburn, 1998), generalized anxiety disorder (Butler et al., 2006; Roth & Fonagy, 1996, 2005; Westen & Morrison, 2001), obsessive-compulsive disorder (van Oppen et al., 1995), panic disorder (Barlow, Gorman, Shear, & Woods, 2000; Butler et al., 2006; DeRubeis & Crits-Christoph, 1998), post traumatic stress disorder (Foa & Rothbaum, 1998), social phobia (Butler et al., 2006), schizophrenia (Pilling, et al., 2002; Zimmermann, Favrod, Trieu, & Pomini, 2005) and a variety of other health problems.

Some meta-analyses have reported that CBT is more efficacious than other types of psychological treatment for depression (DeRubeis & Crits-Cristoph, 1998; DeRubeis, Gelfand, Tang, & Simons, 1999; Dobson, 1989; Hollan, et al., 2005; Pampallona, Bollini, Tibaldi, Kupelnick, & Munizza, 2004), although this finding is not supported by all meta-analyses (Cuijpers, van Straten, Andersson, & van Oppen, 2008; Cuijpers et al., 2010; Gaffan, Tsaousis, & Kemp-Wheeler, 1995; Imel, Malterer, McKay, & Wampold, 2008; Wampold, Minami, Baskin, & Callen Tierney, 2002). There is evidence that CBT for depression can be enduring and may prevent future relapses (Fava et al., 2004; Hollon et al., 2005; Hollan & A. T. Beck, 2004; Scott et al., 2003; Vittengl, Clark, Dunn, & Jarrett, 2007). However, there is concern that some depressed individuals do not respond or only partially respond to CBT (Cuijpers et al., 2008). Therefore there is an ongoing need to identify sound research-based interventions and therapeutic processes to sustain benefits (Friedman & Whisman, 2004; Laurenceau et al., 2007).

Psychotherapy Process

The study of psychotherapy processes has been termed as a 'Rashomon' experience, with limited agreement across populations and across treatments (Mintz, Luborsky, & Auerbach, 1973). Rashomon is the name of a Japanese film in which a crime witnessed by four individuals is described in four mutually contradictory ways. The term 'Rashomon effect' has entered psychological vocabulary to describe the effect of the subjectivity of perception on memory by which observers of an event are able to produce very different, but equally plausible description of that event (Mintz et al., 1973). Psychotherapy processes refer more to how clients and therapists communicate rather than to what they explicity say. They are often preverbal, felt, and intuited whereas content aspects are literal and overt (Schneider, 2001). The varying use of client or therapist or independent observer ratings (Hill & Lambert, 2004; Luborsky, 1994), number of raters (Kazdin, 1977; Shrout & Fleiss, 1979; Tsui, 1983), rater orientation (Raue, Goldfried, & Barkham, 1997), rater qualifications and training (Athey & McIntyre, 1987; Bernadin & Pence, 1980; Shohamy, Gordon, & Kraemer, 1992; Latham, Wexley, & Pursell, 1975; Mercer & Loesch, 1979), rating guidelines and scoring of scale items (Nott, Reeve, & Reeve, 1992), choice of stimulus material (e.g., audiotaped, videotaped or live sessions) (Hill, Nutt, & Jackson, 1994; Hill, & O'Grady, 1985; Horvath & Greenberg, 1994; Jaeger & Busch, 1984; Mercer & Loesch, 1979), unit (e.g., single words, phrases, sentences) to fit the construct (Ekman & Friesen, 1968; Hill et al., 1994), definitions of the same construct (Gaston, Thompson, Gallagher, Cournyer, & Gagnon, 1998; Kiesler, 1973; Llewelyn & Hardy, 2001), sampling methods (i.e., session versus segments or one interview versus several sessions, early, middle or late sessions) (Gaston, et al., 1998; Kiesler, 1973), treatment modalities use of multiple terms (e.g., positive regard, unconditional positive regard, caring and prizing) (Gaston, Marmar, Gallagher, & Thompson, 1991; Gaston et al., 1998; Wolfe & Goldfried, 1988), diagnoses (Hill et al., 1994; Wolfe & Goldfried, 1988), inclusion and exclusion criteria, process and outcome measures, reliability and validity of measures, diagnostic systems, data collection problems (Hill & Lambert, 2004), design sensitivity, data analyses and increasing complexity of theoretical and analytical models may all contribute to lack of consistency within process research (Hill & Lambert, 2004; Llewelyn & Hardy, 2001; Zuroff & Blatt, 2006). Understanding the process to be measured and its theoretical underpinning, making careful methodological decisions and studying the effects of research procedures on the collected data are paramount (Hill & Lambert, 2004).

Alliance Process

In research psychotherapy process is typically linked to outcome and endeavors to demonstrate how therapist-client interactions lead to changes in client symptomology (Hill & Lambert, 2004; Horvath & Symonds, 1991; Pachankis & Goldfried, 2007). Evidence suggests that the alliance between the therapist and the client, as a process, is an important determinant of outcome in psychotherapy, regardless of theoretical orientation (Beutler et al., 2004; Martin, Garske, & Davis, 2000; Horvath & Symonds, 1991; Wampold, 2001). However, the role of the alliance is complex and varies across theoretical approaches (Raue, Castonguay, & Goldfried, 1993). For example, the alliance is considered a curative factor in its own right in psychodynamic (Freud, 1910) and client-centred (Rogers, 1957, 1951) therapies, necessary and important, but not sufficient in CBT (A. T. Beck et al., 1979; J. S. Beck, 1995), and irrelevant in rational emotive behavior therapy (REBT; Ellis, 1996). It is important to understand the theoretical components of the alliance and gain more understanding of this construct in order to measure it accurately within a particular therapy (Andrusyna, Tang, DeRubeis, & Luborsky, 2001; Gaston et al., 1991; Wolfe & Goldfried, 1988). This will ultimately lead to greater clinical benefit.

Alliance Measurement

There are a wide variety of instruments have been developed to quantify the alliance. Currently, there are at least eleven commonly used instruments available that vary in perspective (observer, therapist, or client). Each differs in its theoretical basis and conception of the relationship. It is essential to select an alliance measure that is appropriate to a particular therapy (Safran & Wallner, 1991).

Temporal Relations between Alliance and Outcome

The issue of temporal relations between alliance and outcome has been a controversial and much disputed subject within the field of process research and findings have been inconsistent. It is unclear whether the alliance predicts outcome (e.g., Barber, Connolly, Crits-Christoph, Gladis, & Siqueland, 2000; Klien et al., 2003; Strauss et al., 2006; Zuroff & Blatt, 2006) or, alternatively, whether prior symptom improvement as a consequence of specific therapeutic intervention predicts relational factors (eg., Barber et al., 2000; Safran, & Wallner, 1991; Feeley, DeRubeis, & Gelfand, 1999; DeRubeis & Feeley, 1990). Furthermore, there is the difficulty of determining whether results reflect a casual role of the alliance or simply a correlation between improved symptomology and high ratings of the alliance (R. J. DeRubeis, personal communication, 2nd September 2006; DeRubeis,

Brotman, & Gibbons, 2005; Llewelyn & Hardy, 2001; Stiles, Agnew-Davies, Hardy, Barkham, & Shapiro, 1998). Establishing temporal relations between alliance process and outcome and identifying and testing the specific ways in which alliance produces change are far from resolved (Kazdin, 2005). It is important to determine a research design that will adequately capture temporal relations.

Temporal Design

Pre-post and pre-mid-post research designs provide only a snapshot of the change process (Hayes, Laurenceau, Feldman, Strauss, & Cardaciotto, 2007; Laurenceau et al., 2007). However, the sequencing of events and/or variables is important and assessments of alliance and outcome are needed on multiple occasions in order to evaluate the relationship between them (e.g., unidirectional, bidirectional, concomitant changes) (Kazdin, 2005, 2006). Furthermore, most research focuses on group averages and less emphasis has been placed on the rich information available within individual case studies (Hayes et al., 2007). There is only one known study (Barber et al., 2000) that has indirectly investigated the reciprocal effects of the alliance and outcome. These researchers collected their data at multiple points in time and concluded that the two intricately intertwined variables may amplify each other rapidly and that they might need to be measured every session. Therefore, it was important for the current research to utilize a research design that would enable it to measure the alliance and symptomatic change in every session. Utilizing a single-case research design the current research sought to gain more understanding of what facilitates and what inhibits alliance process within CBT for depression and consequently learn about crucial ingredients that effect change. Investigating the source of symptom improvement within a specific therapy has beneficial implications for training and treatment.

Thesis Outline

This thesis documents the historical, contextual, theoretical and empirical background of the alliance process within CBT for depression, and the research process and outcome (see Figure 1.1). Two studies are presented separately, preceded by the relevant theoretical and empirical background. The main study, named throughout this thesis as the 'current research' investigated the role of alliance and symptomatic change within CBT for depression. The other study, outlined in Chapter 7, was conducted to establish inter-rater reliability for the current research and is named the 'rater reliability study'.

The Depression Outcome Study was conducted at The School of Psychology, Massey University, Albany (2006-2009) to examine the relationship between therapist competence in delivering a guiding model to facilitate reviewing, designing and assigning homework developed by Kazantzis, MacEwan, and Datillio (2005), participants' homework adherence, and symptom and cognitive change within CBT for depression. The current research utilized the archived session data recorded on digital video discs (DVDs) from the Depression Outcome Study to rate the alliance process of the first ten sessions of therapy of the first ten client-therapist dyads in the study. These ten clients completed the Beck Depression Inventory–II (BDI-II; A. T., Beck, Steer, & Brown, 1996) at the beginning of every therapy session and this data was used in the current research to assess symptomatic change.

The current research is reported in nine chapters. The first chapter provides an orientating framework and rationale for the current research.

The second chapter is divided into two sections. The first section reports the origins and the theoretical development of the alliance construct and the second section defines the role of the alliance from different theoretical perspectives.

Chapter 3 begins with a review of empirical evidence demonstrating the relationship between alliance and outcome within the different theoretical orientations and populations. It also reviews literature surrounding the measurement of the alliance and various alliance measures to determine the most appropriate assessment measure for the current research. The Working Alliance Inventory- Short Revised (WAI-SR; Hatcher & Gillaspy, 2006) was chosen as the most appropriate scale to measure the alliance within the current research. Hatcher and Gillaspy (2006) utilized confirmatory factor analysis to develop and validate a client-rated version of the WAI-SR, but to date no one has developed an observer version. Therefore, the current researcher, with the assistance of her supervisor (NK), modified the WAI-SR-C to develop the WAI-SR-O. This development of the WAI-SR-O is more fully explained in Chapter 6.

Chapter 4 critically examines literature surrounding temporal relations between alliance and outcome which highlights the variability in findings within alliance research. It also discusses temporal design to determine the most appropriate method of measuring therapeutic process change. Chapters 2, 3 and 4 provide relevant background information

to conceptualise the alliance within CBT and establish a suitable process measure and research design for the current research.

Chapter 5 outlines the aims and research questions of the current research.

Chapter 6 outlines methodology and is divided into two sections. The first section pertains to the Depression Outcome Study and describes participant characteristics, selection criteria, initial telephone assessment, pre-treatment measure, therapist training and supervision, therapist adherence and competency measures, client suitability measure, outcome measure and treatment received by the participants. The second section of this chapter pertains to the current research and describes the process measure. It outlines the modification of the WAI-SR-C to develop the WAI-SR-O and the guidelines to rate the WAI-SR-O, the characteristics of the raters and the rating of the WAI-SR-O. This chapter concludes with ethical considerations and discusses the single-case research design and data analytic procedures relevant to the current research.

Chapter 7 reviews literature surrounding rating errors and biases, rater training and retraining and characteristics needed by raters for process research. It details rater training for the WAI-SR-O. It also presents the two stages of the rater reliability study. The first stage investigated whether giving more information (i.e., training participants with guidelines developed for rating the WAI-SR-O) would lead to less variance in ratings and higher inter-rater reliability. Stage Two was conducted to test the findings of Stage One and investigated whether there was higher inter-rater reliability rating the alliance within CBT if raters had work experience in the mental health field or a related practice and/or CBT training experience. The methodology and the results of Stage One and Two are presented, followed by a discussion of the results of both stages.

Chapter 8 presents the results of the current research. It is divided into two sections. The first section reviews the inter-rater reliability for the WAI-SR-O in the current research. The second section sets out case results.

Lastly, Chapter 9 presents the discussion of the results, including comments on the implications of the current research for clinical practice, strengths and limitations of the current research and directions for future research.

Psychotherapy
Process
Chapter 1

Depression
Symptomology
Chapter 1

Alliance -
Conceptualization
Chapter 2

Alliance
Measurement -
Chapter 3

Temporal
Relations -
Chapter 4

Aims and
Research
Chapter 5

Training Raters
Rater Reliability
Study -Chapter 7

Methodology
Chapter 6

Stage One -
Evaluation of
Guidelines to
rate
WAI-SR-O/
Results –
Chapter 7

Stage Two-
Evaluation of
Rater Training
& Clinical
Experience/
Results
Chapter 7

Results of
Current
Research
Chapter 8

Discussion –
Current
Research
Chapter 9

Discussion –
Rater Reliability
Study- Stages One
and Two
Chapter 7

Figure 1.1: Thesis Outline

Chapter 2: The Alliance

Overview

The alliance has been proposed as an essential ingredient of psychotherapy (Freud, 1910; Gaston, 1990; Rogers, 1951, 1957; Wampold, 2001). It has been broadly defined as the collaborative, affective bond between the therapist and client (Krupnick et al., 1996). Theorists and practitioners have used various terms to describe different aspects of this relationship, such as, the therapeutic relationship, the therapeutic alliance (Zetzel, 1956), the helping alliance (Luborsky, 1994; Luborsky, Crits-Christoph, Alexander, Margolis, & Cohen, 1983), the therapeutic bond (Orlinsky & Howard, 1986a) and the working alliance (Greenson, 1965). Each term was generated with a particular theory of relationship development in mind. In alliance literature the terms working alliance and therapeutic alliance are often used interchangeably, and despite the popularity and frequency in which the term alliance is used and studied there still appears to be no consistently employed definition. Different researchers often emphasize different aspects (Samstag, 2006). Unless otherwise specified, the general construct under discussion in this thesis will be referred to as 'the alliance'. Horvath and Luborsky (1993) advocated using the term 'the alliance' to describe the therapeutic relationship formed between the therapist and client. Although the role of the alliance varies across therapies, most current theoretical definitions of the alliance have three themes in common: (a) the collaborative nature of the relationship, (b) the affective bond between the client and the therapist, and (c) the client's and therapist's ability to agree on treatment goals and tasks (Bordin, 1979; Gaston, 1990; Horvath & Luborsky, 1993; Horvath & Symonds, 1991).

This chapter reports the origins and the theoretical development of the alliance construct. It defines the role of the alliance from different theoretical perspectives, including psychoanalytic, humanistic (i.e., client-centred therapy, existential therapy, experiential therapy and learning (behavior therapy, rational emotive behaviour therapy, cognitive therapy and cognitive-behaviour therapy) to contrast the role of alliance within CBT with that of other theoretical orientations. The main aim of this chapter is to understand the alliance construct and operationalise it within CBT.

Origins of the Alliance Construct

From the end of the 19th century to around 1960, Sigmund Freud (1856-1939) and his colleagues were dominant influences in psychotherapy (Lambert, Bergin, & Garfield, 2004). Therefore, as with many other aspects of psychotherapy the concept of the alliance owes its genesis to psychoanalysis (Freud, 1910). Freud's (1910) early works explored the difference between the neurotic aspects of the client's attachment to the therapist (i.e., transference). The term transference refers to desires, thoughts, feelings and associated behaviours originating from early (e.g., parental) relationships that are projected or transferred onto the current interpersonal relationship (i.e., with the therapist) (Corey, 2005; Gelso & Carter, 1998). Later he talked about the reality-based collaboration between the therapist and the client, a cojoint effort to conquer the client's pain. This dyadic interaction was maintained by the warm feelings, understanding, and positive regard the therapist had toward the client. This line of thought was extended or modified by other psychoanalytic scholars (e.g., Greenson, 1965; Sterba, 1934; Zetzel, 1956).

Richard Sterba (1934) stated that work was done through the mediation of the ego. He emphasized the client's ability to work in analysis and the need for an alliance between the client's mature ego functioning and the working style of the therapist. The positive identification with the therapist motivated the client to work toward the accomplishment of therapeutic goals. Elizabeth Zetzel (1956) first introduced the alliance as a positive non-neurotic attachment between the therapist and client. She indicated that the therapeutic alliance was separate and distinct from the transferential relationship. Zetzel (1956) maintained that the alliance was formed by the client's attachment and identification with the therapist and that it was necessary in order for the client to withstand the analysis of transference. These early theorists (i.e., Freud, 1910; Sterba, 1934; Zetzel, 1956) emphasized therapist contributions to the alliance.

The next important work came from Ralph Greenson (1965). He introduced the term working alliance and conceptualized the alliance as the client's motivation and ability to work in treatment. This ability to work is fostered by the client (e.g., motivation to overcome the problem, a sense of helplessness, a rational willingness to cooperate and the ability to follow the instructions and insights of the therapist), the therapist (e.g., understanding, insight, empathy, and non judgmental attitude) and the therapeutic interaction. He divided the therapist-client relationship into three components: the transference relationship, the working alliance and the real relationship. These components

were intertwined, but could be readily separated to facilitate a discussion of the analytic process (Greenson, 1965). Greenson (1965) posited that the alliance was neither a technical intervention nor a therapeutic process in itself, but rather a precursor for both.

Pantheoretical Conception of the Alliance

Some researchers have contended that the alliance is pantheoretical; that is, the alliance is important regardless of the type of therapy in which the therapist and client are engaged (Bordin, 1979; Luborsky, 1994; Orlinsky & Howard, 1986). Bordin (1979, 1994) was the first theorist to discuss the alliance outside the context of psychodynamic therapy. His formulation did not clarify the relationship of alliance to transference. He conceptualized a more operational, pantheoretical definition of the alliance consisting of bond, goal and task components. Bond referred to the positive affective attachments that developed between the client and the therapist (e.g., trust, acceptance and confidence) in therapy that allowed the client and therapist to work collaboratively in order to resolve the client's problems. Goal referred to the aims or target outcomes that both the therapist and client agreed would be the most efficacious for the client. Task constituted the activities that that both the therapist and client agreed would be performed to accomplish the goals of therapy. Bordin (1979, 1994) contended that for therapy to be successful both the therapist and client would have to take the responsibility of collaborating on tasks in therapy that would achieve the desired goal. He proposed that the alliance was not in and of itself curative, rather it was a vehicle that allowed change to occur within therapeutic interventions. However, Bordin (1979, 1994) predicted that different therapies would place different demands on the alliance, and that the role of the alliance would be different across therapeutic orientations. Bordin's theoretical formulation has influenced empirical investigations of the alliance and opened the way to quantifying this construct (Horvath & Symonds, 1991).

The Role of Alliance Across Different Treatment Orientations

Within alliance-outcome research there has been little written conceptualizing the alliance construct of the therapeutic modality utilized in the study. However, different psychotherapy systems capture different aspects of human functioning and clinical reality. They focus on and describe different phenomena (Clark, 1995; Gilbert & Leahy, 2007; Lambert & Ogles, 2004; Safran & Segal, 1996) and offer different psychological explanations for various disorders (Wampold, 2007). While the alliance in general may be similar across modalities, governing principles for defining alliance within each theoretical

orientation varies and various specific aspects of the alliance may be emphasized, formulated or used differently (Gaston et al., 1991). To facilitate the alliance construct within CBT and demonstrate its differences to other orientations the following section will briefly outline the role of alliance within the different orientations including: psychoanalysis, client-centred therapy, experiential therapy, existential therapy, behaviour, therapy, rational emotive behaviour therapy, cognitive therapy and CBT.

The Psychoanalytic Perspective

The goal of psychoanalysis is to make the unconscious conscious (Freud, 1910; Gladding, 2007). Within unstructured and non-directive psychoanalytic therapies, the therapeutic relationship between the client and the therapist represents the primary vehicle through which change is assumed to take place (Gaston et al., 1995; Goldfried & Davidson, 1976). It is curative in and of itself and is a necessary condition to provide interpretations to the clients (Gaston et al., 1995). In psychoanalytic therapies it is theoretically anticipated that transference will develop, and that interpretation of the transference is key to the client's therapeutic growth. It is also important for the therapist to focus on their own reactions during therapeutic sessions. By continually observing their own behaviour and emotional reactions the therapist can question what the client may have done to bring about such reaction or conflict (Freud, 1910; Goldfried & Davidson, 1976). A strong affective bond (i.e., warm feelings, understanding, and positive regard) is needed between the client and therapist for this relationship to develop (Been & Winston, 1998; Freud, 1910; Gaston et al., 1995).

The Humanistic Perspective

Client-Centred Therapy

Client-centred therapy, developed by Carl Rogers (1951, 1957) in the 1940s and 1950s, was a significant departure from the Freudian point of view. Rebelling against the 'therapist-as-expert' who interpreted the client's behaviour, Rogers (1951, 1957) contended that the therapeutic relationship formed between the therapist and the client was the only vehicle of change necessary to help clients repair their maladaptive patterns of interaction. He maintained that the alliance or therapeutic relationship, composed of the therapist offered conditions of empathy, unconditional positive regard and congruence was necessary and sufficient for client change to occur (Norcross, 2002; Rogers, 1957). These conditions can be related to John Bowlby's Theory of Attachment, which stems from the need to become attached, not necessarily to get love, but to be more secure by getting

security from others (Bowlby, 1982, 1988). The aim of client–centred therapy is to provide a safe, warm environment to assist clients in their growth process, so that they can better cope with problems they are currently facing, as well as deal with future issues (Rogers, 1951, 1957).

Existential Therapy

Existential therapy is an expression of the humanistic perspective. It evolved from Socratic, Renaissance, Romantic and Asiatic sources, but it was not until the mid-19th century that existential philosophy was formalized from the perspective of Soren Kierkegaard (1844/1980) and M. Heidegger (1926/1962) (Cooper, 2003; Schneider, 2003). In existential therapies therapeutic techniques are secondary to the establishment of a trusting relationship that allows the therapist to challenge the client (Begental & Sterling, 1995; Walsh & McElwain, 2002). The alliance is characterised by mutual respect, individual uniqueness, authenticity, and pursuit of meaning. An agreement or contract is developed between the therapist and client and both contribute and hope to gain from the arrangement (Begental & Sterling, 1995). The therapist and client work through situational and character resistances, explore client-therapist collusion (transference and countertransference) and work with transferential elements and prepare for termination (Begental & McBeath, 1995; Begental & Sterling, 1995). The therapeutic contract is viewed as providing a foundation for therapy with the alliance being conceptualized as a container that will hold the struggles, emotions and relationships necessary to a major life undertaking (Begental & Sterling, 1995). Existential therapists encourage an honest and mutually open relationship, the therapist and client address one another equally, and the therapist strives toward demystification of the therapy process, answering all questions fully and openingly, as opposed to remaining impassive in an effort to evoke transferential distortions (May & Yalom, 2005).

Experiential Therapy

Experiential therapy is another expression of the humanistic perspective. It has its roots embedded in client-centred, existential and gestalt approaches to psychotherapy (Elliott & Greenberg, 2002; Orlinsky & Howard, 1995; Watson & Greenberg, 1998). The alliance is hypothesized to be the core ingredient of change in experiential therapy (Elliott & Greenberg, 2002). However, other processes (e.g., the role of emotional arousal and clients experiencing in the change process) are seen as being equally important in helping clients to resolve specific cognitive-affective problems that brought them to therapy (Watson &

Greenberg, 1998). Experiential therapists emphasize all three components of the alliance including tasks, goals and bonds (Watson & Greenberg, 1998). First, the bond is the therapist's empathetic attunement to the client and the communication of genuine empathy and valuing of the client. The second part involves task and goal collaboration; with the bond this dimension makes up the 'safe working environment'. Thirdly there is the actual work of therapy which involves the use of particular technical interventions (e.g., two-chair work) (Watson & Greenberg, 1994). In practice, experiential therapists strive to maintain a creative tension between the client-centred view of emphasizing a genuine, warm and empathetic alliance and a more active task-orientated, process-directive style of engagement that promotes deeper internal experiencing (Elliot & Greenberg, 1995; Watson, Greenberg, & Lietaer, 1998). If the therapist-client relationship becomes problematic, then it becomes the focus of attention, because a safe working environment is essential for the task of experiencing (Watson & Greenberg, 1998).

Learning Perspectives

<u>Behavioural Therapy</u>

Behaviour therapy became a formal approach to the treatment of psychological disorders in the 1950s and 1960s (Goldfried & Davison, 1976). In contrast to the client-centred tradition in which empathy and warmth are seen as essential in facilitating self acceptance in the client, early behavioural theorists and therapists did not appear to place emphasis on the alliance within therapy and its affect on outcome (Clark, 1995; Gaston et al., 1995; Safran & Segal, 1996). They traditionally viewed technical interventions as the active ingredients of treatment, and the relationship between the therapist and the client was minimized or ignored (Clark, 1995; Persons, 1989). The concept of alliance was hard to define operationally and to measure, and given behavioural therapists commitment to empirical evidence it made sense that they focused their attention on observable therapist and client factors (Sweet, 1994). However, with the development of the social learning perspective behaviour therapists started to pay attention to and study the alliance (e.g., Ford, 1978). They also started to understand that both skills and interpersonal sensitivity were needed on the part of the therapist to form a working partnership (Bandura, 1969). It was found that the positive and supportive attention of the therapist increased the chances of the client being receptive to therapy (Datillio, 2001, 2004; Dattilio, Freeman, & Blue, 1998; Leahy, 2001; Raue et al., 1993; Raue et al., 1997; Safran & Segal, 1996; Staples, Sloane, Whipple, Cristol, &Yorkston, 1975). Alliance in behavior therapy was consequently defined as clients being regarded as co-therapists with decisions about goals

and methods of treatment arrived at mutually. The therapist was looked upon as the consultant, with the eventual aim of treatment being that the clients obtained the skills they needed to be their own behavioral managers (Sweet, 1994). The alliance facilitated between-session risk taking; encouraged clients to carry out various forms of homework; facilitated any modeling; kept up the clients hope and expectancies; and overcame any resistance or noncompliance that may have existed in therapy (Gaston et al., 1995).

Cognitive Therapies

Rational Emotive Behaviour Therapy (REBT)

REBT was founded by Albert Ellis in 1955 (Ellis, 1962, 1973, 1996). Ellis came from a psychoanalytic background, but the core features of REBT have a distinct philosophical emphasis and an existential-humanistic outlook. REBT therapists consider the alliance between the client and the therapist to be irrelevant (Dryden & Ellis, 2001; Ellis & Dryden, 1997). They see themselves as educators and strive to establish the most appropriate learning climate for each client. Favouring a more informal style of delivering therapy, REBT therapists employ a humorous, active and directive, and confrontational style of therapeutic participation (Dryden & Ellis, 200; Ellis & Dryden, 1997). REBT advocates that therapists give unconditional acceptance rather than giving undue warmth and approval to clients. There is a concern that therapist warmth may unwittingly encourage clients to strengthen their need for love and approval. (Dryden & Ellis, 2001; Ellis & Dryden, 1997)

Cognitive Therapy

Cognitive therapy developed in the 1960s as a result of Aaron Beck's research on the psychodynamic theory of depression (Clark & A. T. Beck, 1999). It was here that Beck observed that depressed people had predictable cognitive patterns involving negative views of the self, the world and the future (i.e., the depressive triad) (A. T. Beck, 1976; A. T. Beck et al., 1979; Clark & A. T. Beck, 1999). Cognitive therapy suggests that psychological disorders do not arise from events per se, rather problems arise from the meanings people give to events, filtered through the framework of core beliefs and assumptions that have developed through life experience (A. T. Beck et al., 1979; Padesky, 2004). Cognitive therapy can help clients' correct faulty information processing; modify clients' dysfunctional beliefs that maintain maladaptive behaviours and emotions and provide clients' with skills and experiences that create adaptive thinking (A. T. Beck et al., 1979). In cognitive therapy the alliance serves as a positive background for the actual work

of therapy to be carried out in. Aaron Beck emphasized that therapist characteristics, such as warmth, empathy, respect, genuineness and confidence were necessary in therapy to engage their clients in a process of 'collaborative empiricism' (A. T. Beck et al., 1979). However, he also believed that these non-specific factors were not sufficient to produce optimum therapeutic effect. There is some evidence linking collaboration to better alliance and more positive outcome (Chen & Bernstein, 2000; Kowalik, Schiepek, Kumpf, Roberts, & Elbert, 1997; Tryon & Winograd, 2002).

Collaborative empiricism encapsulates the idea that the client and the therapist work as a team. The client and therapist work together to define problems, set goals, devise adequate experiments to support or refute hypotheses and then adapt the hypotheses based on feedback from therapy interventions (A. T. Beck et al., 1979; J. S. Beck, 1995; Blackburn & Davidson, 1995; Kelly, 1955; Kuyken, Padesky, & Dudley, 2008, 2009; Newman 1998; Padesky & Greenberger, 1995; Raue & Goldfried, 1994). The use of Socratic questioning (i.e., the client arrives at logical conclusions based on the therapist's questions that stimulate curiosity and inquisitiveness) and guided discovery (i.e., a series of inductive questions to reveal dysfunctional thought patterns and behaviour) encourages the client to openly explore and discover their thoughts, feelings and behaviour for themselves (Blackburn & Davidson, 1995; Padesky, 1993; Padesky & Greenberger, 1995; Young, Rygh, Weinberger, & A T. Beck 2008;). A collaborative relationship can be established by providing a clear rationale for the therapy process and each technique that is used, establishing an agenda in a collaborative manner to structure the therapy session, by listening carefully and summarizing regularly and finally by eliciting feedback and providing feedback at the end of each session (A. T. Beck et al., 1979; J. S. Beck, 1995; Blackburn & Davidson, 1995; Padesky & Greenberger, 1995). This active partnership is believed to facilitate the best progress toward the client's goals, teach the client skills to increase self efficacy and help them manage life's problems (i.e., clients are taught new behavioral and cognitive strategies to improve their responses to negative life experiences). The aim of cognitive therapy is to teach the client to become their own therapist in order to prevent relapse (J. S. Beck, 1995; Blackburn & Twaddle, 1996; Padesky & Greenberger, 1995). A bond of trust and positive regard is established and a secure background created in which it was safe to discuss difficult situations, including rifts or ruptures in the alliance. It also provides a safe context for taking risks and trying out new ways of thinking and behaving (A. T. Beck & Weishaar, 1989; Clark, 1995).

Cognitive-Behaviour Therapy

Behaviour therapy and cognitive therapy have merged and the resulting combination, CBT, now dominates clinical research and practice (Rachman, 1996). CBT is based on the assumption that cognitive restructuring will mediate or lead to behavioural changes (Dobson & Dozois, 2001). It recognizes that people's beliefs about the consequences of their behaviours can be key to maintaining unhelpful behaviours (Gilbert & Leahy, 2007; Goldfried, 2003). This is accomplished through relaxation, exposure and a series of directed inquiries into the nature of those beliefs, self-monitoring, problems solving skills training and in-vivo experiments. (A. T. Beck et al., 1979; Blackburn & Davidson, 1995; Padesky, 1993; Padesky & Greenberger, 1995; Persons, 1989).

Much of the governing principles defining the alliance within CBT have developed from both behavioural and cognitive therapies. Similar to cognitive therapy, the CBT therapist shows understanding, appropriate empathy and personal warmth and engages their client in a process of 'collaborative empiricism' to develop a healthier style of thinking, build coping skills and reverse unproductive patterns of behaviour (A. T. Beck at al., 1979; Persons, 1989). Although a solid therapist and client relationship was mentioned by A. T. Beck et al. (1979) they did not emphasize this aspect of therapy (Wright & Davis, 1994). However, in the more recent years CBT has increasingly viewed the alliance as a specific factor in therapy and not just as a vehicle to facilitate the different techniques (Giovazolias, 2004). This view may be due to the widening interest in the scope of cognitive psychotherapy and the growth of work done on information-processing paradigms (Dryden & Trower, 1988). Furthermore, it may have been more influenced by the constructivist approach of people being viewed as change agents who are active creators of their own reality (Giovazolias, 2004). Constructivist therapies view therapist characteristics such as warmth and empathy as being critically important to establish a safe and supportive therapeutic environment (Carmin & Dowd, 1988; Giovazolias, 2004). The role of the alliance within CBT has evolved and continues to evolve through research findings and theoretical developments (e.g., the emerging of cognitive concepts) (Dobson & Dozois, 2001) and practical necessities (e.g., CBT being applied to populations such as Personality Disorders) (A. T. Beck, Freeman, Davis, & Associates, 2004). It is also encouraged by research supporting the reliability and validity of measures of the alliance (Waddington, 2002).

Summary

The alliance in general may be similar across modalities. However, governing principles for defining alliance within each theoretical orientation varies and various specific aspects of the alliance may be emphasized, formulated or used differently. Most theoretical orientations advocate a strong affective bond between the client and therapist and a warm supportive environment. However, REBT therapists prefer to give unconditional acceptance rather than warmth and approval to clients (Dryden & Ellis, 2001). Psychoanalytic and humanistic modalities consider the interactive relationship between the client and the therapist as a central and essential curative element in psychotherapy (Begental & Sterling, 1995; Gaston et al., 1995; Greenson, 1965; Kirschenbaum & Jordan, 2005; Rogers, 1951, 1957; Walsh & McElwain, 2002; Watson & Greenberg, 1998). In contrast, behavioural and cognitive therapies traditionally see a positive alliance as necessary, but not sufficient for therapeutic change to occur (A. T. Beck et al., 1979; Gaston et al., 1995). Behaviour and CBT therapists tend to focus on problems outside the client-therapist relationship. See Table 2.1 for a summary defining alliance across different theoretical orientations Theoretical bases of various therapy modalities suggest that the alliance may be utilized differently or play different roles in therapy. Alliance processes differing across therapies suggests implications for training and supervising new therapists, evaluating competence in the delivery of particular modalities and training raters to rate alliance and for the measurement of the alliance within a specific modality.

CBT therapist characteristics of warmth, empathy and respect are being seen more and more important to establish collaborative relations (i.e., collaborative empiricism) between the therapist and client in order to carry out the goals and tasks of therapy. There has been some research surrounding therapist conditions (e.g., empathy, unconditional positive regard) and whether these conditions are successful in therapy. Rogers (1951, 1957) views about relationship between the therapist and the client are still the focus of much debate (Kirschenbaum & Jordan, 2005; Watson, 2001, 2007). Initial investigations examining the tenability of these therapist conditions as sole agents of change demonstrated that those therapists that provided these conditions were more successful in therapy than those who failed to provide such conditions (Horvath & Luborsky, 1993). Other research evidence refuted these claims (Horvath & Luborsky, 1993; Lambert, de Julio, & Stein, 1978; Rachman & Wilson, 1980). The majority of the findings indicated that it was the client's perception of the therapist as an empathetic individual, rather than actual therapist behaviour that yielded the most robust correlation with outcome (Horvath & Luborsky,

1993). However, while many of Rogers (1951, 1957) tenets were not empirically supported, he did successfully define several important aspects of the alliance that were not previously expressed in the psychodynamic conceptualization of the working relationship (Goldfried, 2007; Watson, 2007). More recently, after an extensive review of the literature, members of the Task Force 29 (i.e., a committee sponsored by the Division 29 of the American Psychological Association to investigate empirically supported therapy relationships) suggested that therapist empathy is essential to psychotherapy and that congruence and acceptance are probably effective (Norcross, 2002). Goldfried (2007) suggested that the while field of psychotherapy has moved beyond Roger's necessary and sufficient conditions as sole agents of change, it recognizes that therapist acceptance is essential and nondirective methods can be effective in improving client motivation.

Within CBT the therapist and client work as a team to set therapy goals and to teach the client skills which the client can then generalize to out-of-session perceptions and behaviours (A. T. Beck et al., 1979; J. S. Beck, 1995; Blackburn & Davidson, 1995; Kelly, 1955; Kuyken et al., 2008, 2009; Newman 1998; Padesky & Greenberger, 1995; Raue & Goldfried, 1994). Conceptualization of the alliance within CBT emphasizes components of Goal, Task and Bond and appears to fit with Bordin's (1979) model of the alliance.

The following chapter discusses measurement of the alliance to determine the most appropriate alliance measure for the current research.

Table 2.1: Summary Defining Alliance Across Different Theoretical Orientations

Theoretical orientation	Therapeutic change	Focus of alliance	Therapist Characteristics
Psychodynamic Therapies	Essential curative element	Focus is on the transferential aspects of the relationship. Strong affective bond between client and therapist.	Warm feelings, understanding, and positive regard.
Client-centred Therapy	Essential curative element	The attitudes of the therapist facilitate change in the client	Empathetic understanding, unconditional positive regard and congruence.
Existential Therapy	Essential curative element	Therapist and client work through situational and character resistances, explore client-therapist collusion (transference and countertransference)	Respect, honesty and a mutually open relationship
Experiential Therapy	Essential curative element	Empathetic attunement. Exploratory empathy. Emphasize all three components of the alliance - Task, Goal and Bond.	Warm and empathetic to promote 'safe working environment'.
Behavioural Therapy	Necessary, but not sufficient	Clients are regarded as co-therapists with decisions about goals and methods of treatment arrived at mutually. Used as a tool to encourage participation within session activity and between session risk taking.	Empathy and warmth.
REBT	Necessary, but not sufficient	Therapists are educators that favour an informal style of delivering therapy. They employ a humorous, active and directive, and confrontational style of therapeutic participation.	Unconditional acceptance only.
Cognitive Therapy/CBT	Necessary, but not sufficient	'Collaborative Empiricism'. Used as a tool to encourage participation within session activity and between session risk taking. Emphasizes all three components of the alliance –Goal, Task and Bond.	Warmth, empathy and respect

Chapter 3: Measurement of the Alliance

Overview

As outlined in Chapter 2, historically there has been a gradual development, though uncoordinated, understanding of the alliance contruct. This chapter focuses on the measurement of the alliance and begins with a review of the empirical evidence demonstrating the relationship between alliance and outcome. Many instruments have been developed to measure the alliance and the four main instruments considered for the current research will be reviewed. The Working Alliance Inventory (WAI; Horvath & Greenberg, 1989), which is explicitly derived from Bordin's theory (1979, 1994), appeared to be the most appropriate instrument to conceptualise alliance within CBT for depression. There have been concerns surrounding whether the alliance should be measure as one general construct or as two independent factors (Andrusyna et al., 2001; Hatcher & Barends, 1996; Hatcher & Gillaspy, 2006; Tracey & Kokotovich, 1989). Therefore, this chapter briefly reviews factor analysis exploration of the subscales of the original 36 item WAI and the two shortened versions of the WAI. The Working Alliance Inventory- Short Revised (WAI-SR; Hatcher & Gillaspy, 2006) was chosen as the most appropriate scale to measure the alliance within the current research. Hatcher and Gillaspy (2006) utilized confirmatory factor analysis to develop and validate a client-rated version of the WAI-SR. Rationale for choosing the WAI-SR over the Working Alliance Inventory- Shortened Version (WAI-S; Tracey & Kokotovich, 1989) is provided. The current researcher, with the assistance of her supervisor (NK), modified the WAI-SR-C to develop the WAI-SR-O. The development of the WAI-SR-O is described further in the methodology section of this thesis.

Empirical Evidence

The alliance appears to be the most frequently studied process of change (Castonguay, Constantino, & Holtforth, 2006; Wampold, 2001) and numerous studies have demonstrated that a positive alliance, as measured in many different ways, from different perspectives (i.e., therapist, client and independent observer) is a predictor of therapy outcome across several modalities of psychotherapy (Horvath & Luborsky, 1993). For example, it has been examined in psychodynamic (e.g., Barber et al., 2000; Barber et al., 2008; Barber et al., 1999; Barber et al., 2001; Carroll, Nich, & Roundsaville, 1997), experiential (e.g., Horvath

& Greenberg, 1989; Watson & Geller, 2005; Watson & McMullen, 2005) behavior and cognitive (e.g., Burns & Nolen-Hoeksema, 1992; Constantino, Arnow, Blasey, & Agras, 2005; DeRubeis, Gelfand, Tang, & Simons, 1999; Hardy et al., 2001; Raue et al., 1997), pharmacotherapy with clinical management (DeRubeis et al., 1999; Wilson et al., 1999; Krupnick et al., 1996), eclectic (Sexton, 1996) and interpersonal therapies (Krupnick et al., 1996). Furthermore, it has also been examined in group (Dies, 2003) and online psychotherapy (Knaevelsrud & Maercker, 2006).

The strength of the alliance-outcome has also been examined across a variety of populations, from elderly depressed clients (e.g., Gaston et al., 1998) to substance use disorders (Barber et al., 2008; Barber et al., 2006; Carroll, Nich, & Ball, 2005; Lebow, Kelly, Knobloch-Fedders, & Moos, 2006; Meier, Barrowclough, & Donmall, 2005), eating disorders (Loeb et al., 2005; Wilson et al., 2002; Wilson et al., 1999), anxiety disorders (Newman, Stiles, Janeck, & Woody, 2006), people with schizophrenia (e.g., Hewitt & Coffey, 2005) and personality disorders (Critchfield & Benjamin, 2006; Smith, Barrett, Benjamin, & Barber, 2006). Therefore, regardless of orientation and population, this facet of psychotherapy process has been demonstrated to be crucial to the process of change.

In a meta-analysis (24 studies published between 1978 and 1990) Horvath and Symonds (1991) revealed a medium-sized effect of .26 between the overall alliance and positive psychotherapy outcome. Martin et al. (2000) found (58 published studies and 21 unpublished doctoral or masters theses between 1977 and 1997) a slightly smaller alliance-outcome correlation, but still in the medium-sized effect range of .22. Results from a meta-analysis (40 studies published between 1990 and 2000) conducted by Beutler et al. (2004) were consistent with previous meta-analyses. They found a medium-sized effect of .22, which supported the alliance being modestly related to treatment outcomes. Although the size of the alliance–outcome relationship was not large, it appeared to be robust. It was also concluded that the construct of the alliance still needed further definition and refinement and that it may be dependent upon how it is measured (Beutler et al., 2004).

Instruments Developed to Measure the Alliance

Subsequent to much interest surrounding alliance–outcome research, a wide variety of instruments have been developed to quantify the alliance. Currently, there are at least eleven commonly used instruments available that vary in perspective (observer, therapist, or client). Each differs in its theoretical basis and conception of the relationship. Until

recently (Cecero, Fenton, Nich, Frankforter, & Carroll, 2001; Fenton, Cecero, Nich, Frankforter, & Carroll, 2001; Safran & Wallner, 1991; Tichenor & Hill, 1989) there has been little information regarding the psychometric properties of these instruments. This thesis will discuss the four most frequently used alliance instruments; The Penn Helping Alliance Scale (Penn; Luborsky, McLellan, Woody, O'Brien, & Auerbach, 1985), The Vanderbilt Therapeutic Alliance Scale (VTAS; Hartley & Strupp, 1983), The California Psychotherapy Alliance Scales (CALPAS; Gaston & Ring, 1992), and The Working Alliance Inventory (WAI; Horvath, 1981; Horvath & Greenberg, 1989). It is essential that the alliance be measured in a way that is appropriate to a particular therapy (Safran & Wallner, 1991).

Penn Helping Alliance Scale

The Penn arose from a psychodynamic perspective and were influenced by Luborsky's work on the alliance (Luborsky et al., 1985). The Penn consists of two types of alliance: Type I (items 1-6) focuses on the client's experience of the therapist as a warm, supportive, and helpful, and Type II (items 7-10) assesses the client's sense of working together with the therapist toward treatment goals (Luborsky et al., 1983). Although the Type I alliance reflects the psychoanalytic focus on the clients affective bond with the therapist, this instrument has been used successfully with CBT and drug counseling (Luborsky et al., 1985). The Type II subscale appears more closely related to Bordin's (1979) pantheoretical definition of the working alliance as a mutual agreement between client and therapist on tasks and goals (Luborsky et al., 1983).

Vanderbilt Therapeutic Alliance Scale

The VTAS is comprised of the 44 items on a 6–point likert-type scale, 14 pertain to the patient, 18 pertain to the therapist and 12 pertain to the patient-therapist interaction (Hartley & Strupp, 1983). Theoretically, the VTAS represents a blend of dynamic and eclectic frameworks (Horvath & Luborsky, 1993). The instrument attributes a successful alliance to the presence or absence of six factors: positive climate, therapist intrusiveness, client resistance or anxiety (Langs, 2004), motivation (Greenson, 1965) and client responsibility (Bordin, 1979).

California Psychotherapy Alliance Scales

The development of the CALPAS scales was influenced by both traditional psychodynamic concepts of the alliance and the subsequent work of Bordin (Gaston & Ring, 1992). The CALPAS consists of 24 likert-type items, each rated on a 7 point scale. It is composed of four alliance scales: (a) patient working capacity (PWC) (items 1-6); (b) patient commitment (PC) (items 7-12); (c) patient-therapist agreement on goals and strategies (WSC) (items 19-24); and (d) therapist's understanding and involvement (TUI) (items 13-18) (Gaston & Marmar, 1994; Marmar & Gaston, 1993; Marmar, Weiss, & Gaston, 1989). The CALPAS is the product of various perspectives, including a focus on the patient's affective bond with the therapist (Freud, 1910), the patients ego capacity for a working alliance (Greenson, 1965; Sterba 1934), mutual agreement on tasks and goals (Bordin, 1979) and the therapists role as an empathetic listener (Rogers, 1957, 1951).

Working Alliance Inventory

Horvath (1981) designed The WAI to capture Bordin's (1979) pantheoretical perspective and measure the strength of the alliance across a variety of theoretical orientations. The WAI consists of 36 items, each rated on a 7-point likert-type scale (where a score of 1 indicates that a good alliance is 'never' present, and 7 that it is 'always' present). The measure has 14 negatively worded and 22 positively worded items. It consists of three subscales (i.e., Goal, Task and Bond). The Goals subscale measures the extent to which a client and therapist agree on the goals of therapy. The Tasks subscale measures the extent to which a client and therapist agree on the interventions that form the substance of therapy. The Bond subscale measures the extent to which a client and therapist possess mutual trust, acceptance and respect (Horvath & Greenberg, 1989). Subscale scores can range from 12 to 84, it can be summed to obtain a total score which ranges from 36 to 252 (Horvath, 1981). Three versions of the WAI are available: a client version, a therapist version and an observer version. Originally developed as a self-report instrument for clients and therapists (Horvath & Greenberg, 1989), the WAI has been transformed to an observer version by replacing the appropriate pronouns (Tichenor & Hill, 1989).

Meta-Analyses of the Alliance Scales

Tichenor and Hill (1989) compared six measures of alliance including: CALPAS, Penn, VTAS, and client, therapist and observer versions of the WAI to investigate the degree to which the measures were correlated and discern if these instruments provided relatively the same information. Data was collected from eight cases of brief (12-20 sessions)

psychotherapy (e.g., psychoanalytic, humanistic or behavioural). Cecero et al. (2001) expanded upon Tichenor and Hill's (1989) study and evaluated the same six alliance measures across three treatments. Data for their study was drawn from psychotherapy sessions (i.e., CBT plus disulfiram, twelve-step facilitation (TSF) plus disulfiram, clinical management plus disulfiram, CBT plus no medication or TSF plus no medication) with 60 participants who were videotaped at week 2 during a clinical trial evaluating treatment for cocaine and alcohol dependence. Despite lack of clear definitional consensus, results from both these studies suggested that the CALPAS, Penn, VTAS and WAI (therapist, client and observer versions) had acceptable levels of internal consistency (i.e., alphas were above .90), reliability, and inter-rater reliability (i.e., intraclass correlation coefficients were all .7 or above). Cecero et al. (2001) found that all the measures were highly related and all seemed to tap a similar construct, whereas Tichenor and Hill (1989) suggested that the Penn may be more of a measure of client capacity than the interaction between the therapist and client. Cecero et al. (2001) also found that reliabilities did vary by treatment condition, which suggested that psychometric properties and alliance–outcome relationships may vary across treatments. However, as pointed out by both sets of researchers, these studies are placed in question when various methodological issues are examined. Small sample sizes (in both studies) could have affected the results and analyses were based on measurement of different sessions and a different number of sessions for each case (i.e., 1, 4, 7, and 11 for 12 session cases or 1, 4, 7, and 15 for 20 session cases) or only on a single, early (Session 2) session per case (Tichenor & Hill, 1989 and Cecero et al., 2001 respectively). It is not unreasonable to posit that the alliance may fluctuate across even a single session, but any sample of one session may not be an accurate reflection of the relationship as a whole. Furthermore, to protect client confidentiality, Cecero et al. (2001) designed their study so that only the therapists were visible on the tapes, although both clients and therapists voices were recorded. This limitation of not being able to see the client may have impacted the validity of these ratings (Cecero et al., 2001).

Safran and Wallner (1991) compared the predictive validity of client–rated versions of the WAI and CALPAS. Their results indicated that both instruments were predictive of some outcome measures in short-term cognitive therapy for depression. Fenton et al. (2001) evaluated the predictive validity of the Penn, VTAS, CALPAS and WAI utilizing various rating perspectives (observer-, therapist- and client-rated versions) across two types of treatment (ie., CBT and TSF). They found that all observer-rated instruments were significantly correlated with outcome; whereas client-rated and therapist-rated instruments

did not predict outcome. They suggested that the weak relationship between client-rated and therapist-rated measures may have been due to a highly objective outcome measure (i.e., primary outcome measure was maximum consecutive days abstinent from cocaine while in treatment). Fenton et al. (2001) also reported concerns of inter-rater reliability estimates only being based on a sample of eight sessions and while they held recalibration meetings to correct rater drift, it was possible that drift had occurred.

None of the abovementioned studies reported clinical sensitivity of the alliance measures or whether these measures can capture alliance ruptures and repairs. The alliance can vary across therapy and shifts in the alliance can influence outcomes (Gelso & Carter, 1994; Safran, Muran, Samstag, & Stevens, 2002). Alliance ruptures are an impairment or fluctuation in the quality of the relationship between the therapist and the client (Safran, Crocker, McMain & Murray, 1998). These ruptures may vary in intensity, duration and frequency depending on the therapist-client dyad (Safran & Muran, 2000) and phase of treatment (Cashdan, 1973; Gelso & Carter, 1994; Tracey, 1993). While the current research did not aim to examine the disruption and repair of alliance across therapy or its relationship to outcome, it is important to acknowledge that the working through and repairing of ruptures are an intrinsic part of the change process (Bordin, 1994) and is required to enhance the efficacy of CBT (Whisman, 1993). In his later work Bordin (1994) viewed the process of alliance as consisting of three stages: the initial formulation where the alliance is based on initial impressions, trust and liking, the period of work and then the period of completion. During the period of work Bordon (1994) postulated a process of wear and tear. As a result of the client's problems creating a strain on the alliance this process may weaken the alliance or provide the opportunity for the dyad to work out some issues, thus strengthening the alliance. The WAI allows broad measurement of Goals, Tasks and Bond dimensions on a session by session basis. It does not capture the moment by moment interaction between the therapist and client (Tichenor & Hill, 1989). The CALPAS measures therapist and client positive and negative contributions (Marmar & Gaston, 1993) and more recently it has been utilized to rate "rupture–repair" episodes within cognitive therapy for avoidant and obsessive-compulsive personality disorders (Strauss et al., 2006).

Utilization of the Working Alliance Inventory in the Current Research

The WAI is a popular research measure. It has been translated into several languages (eg., Mandarin, French, Portuguese, Spanish, and German) and has been adopted to some

specific populations (e.g., supervisors and case managers) (Horvath, 2006). The WAI has adequately established alliance in online studies (Cook & Doyle, 2002). The WAI has been utilized in Australia to demonstrate the value of the alliance in case management (Howgego, Yellowlees, Owen, Meldrum, & Dark, 2003). A literature search of MedLine, and PsychInfo databases (between 1990 and 2007) for studies utilizing the WAI to examine the relationship between alliance and outcome did not uncover any published research in New Zealand. Most alliance research has been conducted overseas in North America and the United Kingdom.

The CALPAS, Penn, VTAS, and WAI have significant overlap at the global level which suggests growing agreement the construct exists and that each measure is tapping the same general phenomenon (Bachelor, 1991; Cereroc et al., 2001; Fenton et al., 2001; Hatcher & Barends, 1996; Stiles et al., 2002; Tichenor & Hill, 1989). However, the instruments tend to diverge at the subscale level. The WAI, CALPAS, Penn and VTAS measures are related, but do not have identical, underlying concepts (Horvath & Luborsky, 1993). The following constructs monitored by the abovementioned alliance instruments include: therapists and clients positive and negative contributions (e.g., CALPAS), shared or agreed goals for the therapy (e.g., CALPAS, WAI, Penn), capacity to form a relationship (e.g., Penn, VTAS, CALPAS), acceptance or endorsement of therapy tasks (e.g., CALPAS, WAI) and active participation in therapy (e.g., CALPAS, Penn, VTAS) (Horvath & Luborsky, 1993).

Theoretical diversity is also reflected in the insruments. Some measures, such as the Penn and VTAS were initially developed for psychoanalytic therapy and are therefore less valid for research within CBT (Castonguay, Goldfried, Wiser, Raue, & Hayes, 1996; Safran & Wallner, 1991). Although the Type II subscale Penn may be closely related to Bordin's (1979) pantheoretical definition of the working alliance. The CALPAS (and the VTAS to some extent) represents a more eclectic position and appears to have more independent subscales (Horvath & Symonds, 1991). The WAI appears to be the most transtheoretical of the instruments (Safran & Wallner, 1991) and is more consistent with the format and goals of CBT (Strauss et al., 2006).

Summary

In conclusion, the WAI which is explicitly derived from Bordin's theory (1979, 1994) appeared to be the most appropriate instrument to conceptualise alliance within CBT for depression. It is pantheoretical, has sound psychometric properties and possesses broad clinical utility. It is important to understand the alliance construct and its theoretical underpinnings, and to choose the best instrument to endorse CBTs philosophy. Table 3.1 displays a summery of alliance measure comparisons.

Table 3.1: Summary of Alliance Measure Comparisons

Alliance Measure	Orientation Base	Constructs	Psychometric Properties	Appropriateness for Research in CBT
Penn	Psychodynamic	Shared or agreed goals for therapy. Capacity to form a relationship. Active participation in therapy.	Sound psychometric properties.	Low
VTAS	Psychodynamic	Capacity to form a relationship. Active participation in therapy.	Sound psychometric properties.	Low
CALPAS	Psychodynamic However, mutual agreement on tasks and goals scale based on Bordin's (1979, 1994) perspective.	Therapists and clients positive and negative contributions. Shared or agreed goals for therapy. Capacity to form a relationship. Acceptance of therapy tasks. Active participation in therapy.	Sound psychometric properties.	Moderate
WAI	Based Bordins' (1979, 1994) transtheoretical theory.	Shared or agreed goals for the therapy. Acceptance of therapy tasks. Consists of three subscales (i.e., Bond, Tasks and Goals)	Sound psychometric properties.	Great

Factor Analysis Exploration of the Working Alliance Inventory

More recently literature has pointed out concerns surrounding whether the alliance should be measured as one general factor or as two independent factors in CBT (Andrusyna et al., 2001). Bordin's (1979) model, as measured by the WAI suggested several alliance components. However, researchers have continued to assume a one-factor construct (Hatcher & Barends, 1996). Andrusyna et al. (2001) argued that this assumption could lead to research problems, especially if the alliance is more complex than many first believed. Horvath (1994b) posited that alliance scales mirror the diversity of the theoretical orientation and that there were no theoretical guidelines as to the relative importance of the alliance components. All the scales assume that the components are of equal importance (i.e., weighted equally) and are additive. Andrusyna et al. (2001) speculated that the components of alliance may need to be weighted differently, or explored as independent constructs. Research findings on this matter appeared to be inconsistent. Horvath and Greenberg (1986) reported that the inter-scale correlations were high and therefore their distinctness was questionable. However, Horvath and Marx (1990) demonstrated that the subscales fluctuated independently over time within individual clients, and therefore seemed to be tracking different phenomena in the therapy.

Using confirmatory factor analysis, Tracey and Kokotovic (1989) examined the factor structure of the 36-item WAI within psychodynamic therapy. They picked the four highest loading items on each of the three dimensions to form a 12-item short form of the WAI. They concluded that the therapist and client rating of the 12-item shortened version of the WAI (WAI-T-S and WAI-C-S respectively) appeared to measure primarily a general alliance factor and secondarily the specific aspects of the alliance (i.e., Goal, Task, and Bond). They concluded that the alliance may be composed of both a general higher order multifaceted construct and several more specific aspects or components. Since its development the WAI-S has been used widely in psychotherapy research (Hatcher & Gillaspy, 2006). When Tyron and Kane (1995) investigated the relationship of therapist rated client and therapist involvement and client relatedness in the first session to the strength of the working alliance measured after the third session, and type of client termination (10 therapists of varied orientation and 109 college student clients) they reported high internal consistency estimates for total WAI-S scores (i.e., .92 for the client version and .90 for the therapist version). Utilising hierarchal linear modeling Busseri and Tyler (2003) compared the 36-item (WAI) ratings to the ratings comprising the 12-item scale and found that WAI and WAI-S scores were highly correlated and had comparable

descriptive statistics, internal consistencies and subscale intercorrelations within and across rater perspectives. Predictive validity estimates for the total scales of the WAI and WAI-S were also very similar within rater perspective. These results support the interchangeability of scores on the WAI and WAI-S scales (Busseri &Tyler, 2003).

Hatcher and Barends (1996) investigated three alliance scales, also within psychodynamic therapy and found that the alliance, as measured by the 36-item WAI-C and WAI-T, has two independent factors with Goal and Task items grouping on one factor and Bond items grouping on the other factor. They suggested that the Bond might be strong even if other aspects of the alliance (i.e., Goal and Task) are not established. Trust and mutual liking between the therapist and client may exist even if they have not reached agreement on therapy Goals.

More recently, Andrusyna et al. (2001) reported an exploratory factor analysis of the observer version of the shortened WAI (WAI-S-O) within CBT for depression with 70 patients. They also found that the WAI-S-O consisted of two independent factors. The first factor (agreement/confidence) consisted of all the Goal (e.g., 'The client and therapist are working on mutually agreed-upon goals') and Task (e.g., 'There is agreement on what is important for the client to work on') items and the confidence Bond item (i.e., 'The client feels confident in the therapists ability to help the client'). Andrusyna et al. (2001) postulated that confidence in the therapist's ability to help the client may refer to helping the client with tasks and also helping the client eventually achieve certain goals, rather than to the interpersonal relationship with the therapist The second factor (relationship) consisted of the three remaining Bond (e.g., 'There is a mutual liking between the client and the therapist') items and included those items that relate to the interpersonal relationship between the therapist and the client. Andrusyna et al. (2001) reported that inter-rater reliability of the two raters, estimated by a Pearson correlation coefficient was 0.67. Item–by-item inter-rater reliabilities ranged from a low of 0.14 to a high of 0.65, with a medium of 0.42. These inter-rater and item-by-item inter-rater reliabilities appear to be fairly typical of observer alliance scales (Cecero et al., 2001; Tichenor & Hill, 1989). However, this study was limited due to the use of audiotaped archived data (archived data from the depression study of Jacobson et al., 1996) which did not allow nonverbal behaviours to be noted, and that the analyses were only based on a single early session (i e., Session 2). Andrusyna et al. (2001) also suggested that their study be replicated with the longer version of the WAI in order to have a more precise conceptualisation of the alliance

within CBT, and allow the opportunity for more than one confidence item to be loaded onto the agreement/confidence factor rather than the relationship factor.

Utilising the 36-item client version of the WAI scale, Rector, Zuroff, and Segal (1999) reported significant correlations between the WAI-Goal and WAI-Task during cognitive therapy for a depressive disorder, or anxiety disorder, or a combination of both depressive and anxiety disorders. The WAI-Bond and WAI-Goal, and WAI-Bond and WAI-Task were moderately correlated.They found that the perceived agreement between the client and therapist on the goals and tasks of therapy predicted change in dysfunctional beliefs in CBT, but the bond between the client and therapist did not. However, a strong therapeutic bond seemed to provide the context in which changes in dysfunctional beliefs resulted in significant symptom change (Rector et al., 1999). Rector et al., (1999) suggested that therapist warmth and empathy may help to correct the depressed client's negative cognitive distortions about relationships. Furthermore, a strong bond may facilitate greater disclosure from the client and more active collaboration between the therapist and the client. However, these authors have acknowledged the limitations of their study which included: the study being correlational so therefore it could not determine causality, the alliance was only assessed at pre and post treatment and, as such, bidirectional relationships between alliance and cognitive change could not be assessed and the absence of direct examination of the adherence or competence of specific interventions employed.

Development of the Working Alliance Inventory– Short Revised (WAI-SR)

Hatcher and Gillaspy (2006) used confirmatory factor analysis to develop and validate a revised client-rated short version of the original WAI (WAI-SR), utilising a much larger data base than Tracey and Kokotovic (1989).The first sample consisted of 231 participants receiving psychodynamic therapy. The participants completed alliance measures (i.e., WAI, Penn, HAQ and CALPAS) at different points during their therapy ranging from the 2nd to the 274th session. The second sample was composed of 235 participants receiving treatment approaches of CBT, psychodynamic, person-centered, systemic and "other" interventions. The participants completed the WAI after their third session of therapy. More sophisticated statistical procedures were utilized to derive a parsimonious subset of the full scale WAI instrument to closely parallel the total alliance level obtained in the regular WAI (Hatcher & Gillaspy, 2006). Evidence suggested that the WAI-SR better differentiated Goal, Task, and Bond dimensions, particularly the Task dimension. All

subscale score alphas of the WAI-SR ranged from .85 to .90, and total score alphas were .91 and .92. Of note, one Task item and one Goal item crossed over to load on the other factor. It was possible that clients saw "agreeing on what is important to work on" as referring to Goals rather than Tasks, and being "clearer as to how one might be able to change" as a Task-focused item rather than a Goal-focused item. Also of note, is that the confidence item is no longer part of the Bond dimension and is replaced with "(therapists name) and I respect each other" and "(therapists name) and I mutually trust each other" is replaced with "I feel (therapists name) cares about me, even when I do things that he/she does not approve of" (Hatcher & Gillaspy, 2006). This shift of items has been supported by A. O. Horvath (personal communication, November 5, 2005 between A. O. Horvath and Hatcher & Gillaspy, 2006). Hatcher and Gillaspy (2006) contended that the original 7-point scaling was not optimal for the WAI because clients did not appear to discriminate effectively at the lower ends on the scale, so was replaced by a 5-point scale. Based on Rasch analysis of the scale, adjustments involved combining the bottom three points and middle two points, which improved the quality of the scale (Hatcher & Gillaspy, 2006). Hatcher and Gillaspy (2006) also examined whether the negative items provided useful additional information and found that they did not relate to Bordin's theory. While maintaining that the results of their study helped the theory based WAI to better operationalise the alliance concept, Hatcher and Gillaspy (2006) also acknowledged the limitations of their study, which included the client population being predominantly white and the treatment involved primary dynamic-interpersonal in the first sample and quite diverse treatments in the second sample. The WAI-SR is recommended by A. O. Horvath (2006) to researchers requiring a briefer assessment tool.

Table 3.2: Summary of the Different Versions of the Working Alliance Inventory

Working Alliance Inventory, (36 items) WAI.	Working Alliance Inventory – Shortened Version (12 items) WAI-S.	Working Alliance Inventory – Short Revised Version, (12 items) WAI-SR
WAI –Therapist Version (WAI-T) (Horvath & Greenberg, 1989).	WAI-S-Therapist Version (WAI-S-T) (Tracey & Kokotovic, 1989). Four highest loading items on each of the three dimensions to form a 12-item short form of the WAI.	
WAI –Client Version (WAI-C) (Horvath & Greenberg, 1989)	WAI-S-Client Version (WAI-S-C (Tracey & Kokotovic, 1989). Four highest loading items on each of the three dimensions to form a 12-item short form of the WAI.	Hatcher and Gillaspy (2006) utilised confirmatory factor analysis to develop and validate a revised client-rated short version of the WAI (WAI-SR-C). WAI-SR-C consists of 2 Independent factors. Some items differ to the WAI-S. One Task item and one Goal item crossed over to load on the other factor. A 5-point scale and rephrasing negative items to positive ones was utilized.
WAI –Observer Version (WAI-O) WAI was changed to an observer version by replacing the appropriate pronouns (Tichenor & Hill, 1989).	WAI-S–Observer Version (WAI-S-O) Andrusyna et al. (2001) reported an exploratory factor analysis of the observer version of the WAI-S	The current research modified the WAI-SR-C to develop the WAI-SR-O by matching corresponding items from the original observer 36-item version of the WAI (This modification is discussed further in Chapter 6)

Alliance Measurement Conclusions

Literature to date has suggested that selecting an observer-version of the WAI-SR better operationalises the alliance concept within CBT for the current research. See Table 3.2 for a summary of the different versions of the Working Alliance Inventory. Generally short forms of the WAI have not been criticized for lack of reliability or loss of information. However, Hatcher (2006) questioned whether certain key alliance features, such as being able to adequately capture collaborative processes between the therapist and client, may be beyond the capacity of a brief general alliance measure and whether this would limit the

understanding of the role of the alliance within CBT. Therefore, the current research sought out guidelines from Raue and colleagues (1991) to rate the WAI and then extended and modified these guidelines to rate the WAI-SR-O within CBT. Further information about the development of the guidelines can be found in Chapter 6 and training raters to rate the WAI-SR-O with the guidelines is described in Chapter 7.

In the current research the WAI-SR-O was chosen over the WAI-S-O for the following reasons: the WAI-SR had been recommended by A. O. Horvath (2006) to researchers requiring a briefer assessment tool, Hatcher and Gillaspy, (2006) replaced the original 7-point scaling with a 5-point scale, which is more optimal for the WAI-SR and it has no negative items, and evidence suggested that the WAI-SR-C better differentiated Goal, Task, and Bond components of the alliance. When Andrusyna et al. (2001) conducted an exploratory factor analysis with the WAI-S-O within CBT for depression, they utilised audiotaped archived data and analyses based on a single early session. However, while Hatcher and Gillaspy's study was not without its limitations, their samples were larger, they utilized more sophisticated statistical procedures and sessions were measured at different points of the therapy. Finally the original WAI was successfully transformed into an observer version by Tichenor and Hill (1989). Therefore it was possible that the WAI-SR-C could also be transformed into the WAI-SR-O by matching corresponding items from the observer 36-item version of the WAI. Developing the WAI-SR-O will be discussed more fully in the methodology section of the current research.

Although a clear conclusion about whether or not the WAI does indeed measure three distinct components of the alliance cannot be drawn, there is evidence that the subscales measure different but related constructs and that there may be both a general alliance factor and specific factors measured by the WAI. (Andrusyna et al. 2001; Hatcher & Barends, 1996; Tracey & Kokotovic, 1989). Therefore, it is appropriate to investigate the roles of the alliance subscales (Goal, Task and Bond) within the current research.

To complete this section on alliance measurement Chapter 4 critically examines literature surrounding relations between the alliance and outcome and temporal design to capture therapeutic process change.

Chapter 4: Temporal Relationship between Alliance and Outcome

Overview

The current research project was interested in understanding the role of the alliance and outcome within CBT for depression. The aim of this chapter was to critically examine literature surrounding temporal relations between alliance and outcome. Literature pertaining to this area of research is huge and rather than presenting a review of all published research this chapter will instead focus on key articles appearing from the early 1990s. The chapter concludes with information for determining an appropriate temporal design for the current research.

Alliance–Outcome Research

To date research findings have been inconsistent. Early studies measured the alliance in the midst of treatment and correlated it with symptom change from the beginning to the end of treatment (Castonguay et al., 1996; Gaston et al., 1998; Krupnick et al., 1996; Safran & Wallner, 1991) or they averaged the alliance across the duration of the treatment and then this averaged measure was related to overall symptom change (Gaston et al., 1998; Krupnick et al., 1996). Feeley et al. (1999) proposed that in the aforementioned studies the outcome variable would have included symptom change that had occurred before the alliance was assessed, thus confounding prior and subsequent change. However, there was no way of avoiding this temporal confound when the alliance was averaged across the course of therapy. They suggested that if the alliance was assessed in a given session this temporal confound could be avoided by assessing the symptom change that occurs prior and subsequent to that session. Zuroff and Blatt (2006) proposed that currently, there were two types of evidence that can determine whether the alliance predicts outcome. First, alliance measures taken early in treatment, when therapeutic change presumably has not yet taken place, are significantly related to outcome. Secondly, evidence is provided by studies in which early change in symptoms is statistically controlled and measures of alliance continued to predict change (e.g., Barber et al., 2000; Klein et al., 2003).

Several studies have reported that early change in symptoms predicted a subsequent increase in the alliance (eg., Barber et al., 2000; DeRubeis & Feeley, 1990; Feeley et al., 1999; Safran, & Wallner, 1991). In two studies examining observers' ratings of the alliance in cognitive therapy (e.g., DeRubeis & Feeley, 1990; Feeley, et al., 1999) symptom change was not predicted from alliance scores early in cognitive therapy. The alliance score in late therapy was predicted from the amount of benefit the client had achieved until that point, although this finding only approached significance. The alliance acted like the result of, rather than the cause of, positive therapeutic change. However, the study utilized the Penn which was developed from a psychodynamic perspective. This measure may not have captured the alliance construct within cognitive therapy.

Tang and DeRubeis (1999) found that sudden, substantial and stable improvements in symptoms or 'sudden gains' during the course of cognitive therapy for depression were not predictable from the quality of the alliance just before the sudden gain. They proposed that the quality of the alliance was reliably higher in the session that followed a sudden gain suggesting that positive alliance followed cognitive change. However, Busch, Kanter, Landes and Kohlenberg (2006) pointed out that there have been no analyses of sudden gains over the entire course of CBT for depression, so evidence of this proposal is limited. Tang and DeRubeis (1999) did not investigate pretreatment gains or sudden gains that occurred directly after the pretreatment assessment, but before the first session and first session gains or sudden gains that occurred between the first and the second session. Tang and DeRubeis (1999) excluded first sessions from their study because they deemed them to different from other sessions. However, these sessions may be particularly relevant for evaluation of early response to therapy and/or early alliance (Busch et al., 2006). Plotnicov (1990) maintained that it was quite obvious from everyday clinical experience that first impressions, especially regarding the quality of the alliance may determine whether the client will return for a second visit. Busch et al. (2006) found that positive outcomes were predicted by first session gains and sudden gains occurring in the first half of therapy. These authors suggested that no case could be made for specific therapy interventions influencing these first session gains and that non-specific factors, such as the alliance was therefore potentially important to outcome within CBT for depression. Findings of the abovementioned studies reporting symptom improvement subsequent to an increase in alliance have not been replicated in more recent, larger, better controlled studies (Zuroff & Blatt, 2006).

Strongest evidence for a causal effect of the alliance on outcome are the results from Klien et al. (2003), within a sample 367 individuals with depression, during cognitive-behavioural analysis system of psychotherapy (CBASP) with or without medication for the treatment of depression. While controlling for prior severity and improvement in depression and eight prognostically relevant patient characteristics, these researchers found that early alliance (as measured by client ratings of the WAI at sessions 3-4, 8-12, and 16-20) significantly predicted improvement in depressive symptoms over the course of treatment, but early improvement in symptoms did not predict subsequent alliance. These authors suggested that the effect of symptom reduction on the alliance did not appear until later in the course of therapy. However, the alliance was assessed at only three time points and the timings of these ratings may not have reflected the alliance accurately. Dyads with lower alliances have usually terminated by the third session or later (Tryon & Kane, 1995). Though Klien et al. (2003) argued that differences between clients who were and who were not included in the analyses, due to early dropout or failure to complete assessments, were minor and that they probably did not have a substantial impact of the findings.

Past results that control for early change have not been consistent. In a large sample of cocaine dependent clients receiving several different treatments Barber et al., (1999) reported that client ratings of the alliance, as measured by the client and therapist ratings of the CALPAS at the end of sessions 2 and 5, were not associated with subsequent change on drug related measures after controlling for prior improvement. Therapist ratings were even less predictive. However, in a sample of 86 clients with chronic depression, generalized anxiety disorder, or avoidant or obsessive-compulsive personality disorder treated with supportive-expressive dynamic psychotherapy, Barber et al. (2000) demonstrated that the alliance (as measured by client ratings of the CALPAS at sessions 2, 5 and 10 and then after every fifth additional session) predicted subsequent change in depressive symptoms when prior change in depression was partialed out. Their results indicated that greater symptomatic improvement through session 5-10 led to higher levels of alliance, and that, in turn, alliance was associated with further symptomatic improvement. This suggested that that there may be reciprocal effects between change in the alliance and change in depressive symptoms which needed further investigation. These researchers concluded that the alliance and outcome may be so intricately intertwined that they may amplify each other rapidly and therefore they needed to be measured every session. Whereas, in a sample of 54 elderly depressed clients receiving behavioural, cognitive or brief dynamic therapy Gaston et al. (1991) found that the association between

the alliance (as measured by CALPAS after sessions 5, 10 and 15) and subsequent change was not significant after controlling for improvement in depressive symptoms prior to the assessment of the alliance.

Recent studies have had more positive results. Zuroff and Blatt (2006) used data from the National Institute of Mental Health Treatment (NIMH) for Depression Collaborative Research Program (TDCRP) and early VTAS scores (i.e., those from the third treatment session) from observer raters, to examine the impact on treatment outcome of the patients perception of the quality of the therapeutic relationship and contribution to the therapeutic alliance. They discovered that, across CBT, interpersonal therapy (IPT), and the two medication with clinical management conditions, the perceived quality of the early alliance, adjusted for early clinical improvement, predicted a reduction in depressive symptoms subsequent to the measure of the alliance.

In an investigation of the alliance in cognitive therapy for personality disorders Strauss et al. (2006) used patient ratings of CALPAS scores to assess overall alliance strength at sessions 2, 5, 10, 20, 30, 40, 50, and 52. Treatment consisted of up to 52 weekly sessions. They found that a stronger early alliance predicted more improvement in symptoms of personality disorder and depression, even when the number of sessions completed and early change in depression were statistically controlled. It is possible that an early alliance is particularly important for difficult–to-treat individuals, especially those with personality disorders, as hypothesized by Aaron Beck (A. T. Beck et al., 2004). A strong early alliance can instill hope and provide a solid foundation for the course of therapy (Horvath & Luborsky, 1993; Gaston, 1990).

Other diagnoses had mixed results. Wilson et al. (2002) found that client self-report of the alliance (as measured by The Helping Relationship Questionnaire, HRQ; Luborsky et al., 1996, at the end of session 4 and midtreatment) did not mediate treatment outcome during a study of CBT for bulimia nervosa. Wilson et al. (1999) also demonstrated that the alliance (as measured by HRQ scores obtained at sessions 5, 12, and at termination during CBT and Supportive Psychotherapy treatment) did not predict end-of-treatment binge eating and purging frequencies but that it did predict remission status. Individuals with bulimia nervosa usually report not being close to their families and often have a conflictual relationship with them. They describe their parents as being neglectful and rejecting (Sadock & Sadock, 2003), thus it is hardly surprising that they may find it difficult to form

a strong alliance with their therapist. However, Treasure et al. (1999) reported that the alliance (as measured by client and therapist ratings of the WAI at week 4, during sequential treatment of CBT and Motivational Enhancement Therapy) was positively associated with change in binge eating and purging frequencies in CBT. Loeb et al. (2005) found that early alliance (as measured by observer ratings of the VTAS at sessions 6, 12, and 18, during CBT and Interpersonal Therapy) predicted post-treatment purging frequency in CBT.

Specific diagnoses and varying nature of the disorder may impact alliance ratings (Barber et al., 2000; Klien et al., 2003). Dunn Morrison and Bentall (2006) demonstrated that alliance may not have a direct effect on outcome during cognitive therapy in patients with positive symptoms of schizophrenia. Furthermore, Barber et al. (2006) found that the relationship between therapist adherence to models of treatment and outcome may need to take into account the affect of the alliance when treating cocaine addiction. These authors suggested that the alliance may have a mediating or moderating effect on outcome.

Determining an Appropriate Temporal Design

It was important for the current research to select an appropriate temporal deign to understand temporal relations between the alliance and symptomatic change within CBT for depression. It is commonly thought that change is gradual and linear within psychotherapy research and research designs and statistical analyses utilized to study therapeutic change often reflect that assumption (Hayes et al., 2007). Simple pre-post outcome designs that measure predictors of change once or twice and then correlate these measures with symptom change at the end of treatment only provide snapshots of the change process. They do not adequately capture the slope of change or the mediators (i.e., a mechanism or process that explains why the intervention works) and/or moderators (i.e., a variable that influences either the direction or the strength between an independent variable and dependent variable) of change (Collins & Graham, 2002; Hayes et al., 2007; Kazdin, 2006; Laurenceau et al., 2007). To date process research has been largely limited to pre-mid-post measures of change. However, if the client's symptoms improve during the middle of treatment these improvements may come before change in the mediator. Furthermore, a reduction in symptoms may occur at the same time as the change in the mediator and/or both may have changed due to some other influence (Kazdin, 2006; Laurenceau et al., 2007). Laurenceau et al. (2007) argued for the use of more frequent assessments. Multiple assessments during treatment can be costly (Collins & Graham,

2002), but provide information on the sequencing of processes and outcomes and the possibility of bidirectional changes. It is possible that each one influences the other in some way and at different points in therapy (Kazdin, 2006). Laurenceau et al. (2007) also maintained that the use of statistical techniques such as growth curve modeling, growth mixture modeling and dynamical systems modeling are more capable of determining the mechanisms underlying change than traditional methods (e.g., correlational and ANOVA-based approaches). Moreover, they tend to address nonlinear change in real world context of therapy and capture the fluctuations and dynamics of change (Hayes et al., 2007). For example, Hayes et al. (2007) maintained that patterns of depression severity change, such as, early rapid response (Ilardi & Craighead, 1994), sudden gains (Tang & DeRubeis, 1999; Tang, DeRubeis, Beberman, & Pham, 2005) and depression spike (Hayes, Beevers, Feldman, Laurenceau, & Perlman, 2005) are discontinuous and non linear and would not have been apparent in pre-post analyses of data. Hayes et al. (2007) suggested that these new statistical techniques could be looked upon as a revival of the single–case research design with individual data (Barlow & Hersen, 1984; Blampied, 1999, 2001; Hersen & Barlow, 1976; Kazdin, 1998, 2001, 2003).

Early versus Late Alliance Measurement

It is important to select an appropriate timing of assessments in order to capture the effects of change (Laurenceau et al., 2007). It has been demonstrated that early alliance is a stronger predictor of outcome than middle and late alliance (Constantino, Castonguay, & Schut, 2002; Horvath 1994a). Establishing a quick working alliance may be particularly important in brief therapies (Kokotovic & Tracey, 1990). Plotnicov (1990) reported a slight but definite increase of alliance in the early sessions of therapy, but that it was not strongly correlated with outcome. However, other earlier studies and more recent empirical research have demonstrated that if alliance forms early in treatment it is predictive of later positive therapeutic outcome (e.g., Hartley & Strupp. 1983; Henry & Strupp, 1994; Hersoug, Monsen, Havik, Odd, & Hoglend, 2002; Horvath, 2000; Horvath & Luborsky, 1993; Loeb et al., 2006; Kokotovic & Tracey, 1990; Marzali, Munroe-Blum, & McCleary, 1999; Reandeau & Wampold, 1991; Strauss et al., 2006). A strong early alliance may influence outcome by increasing treatment engagement, instilling hope and expectation of speedy relief, and providing a solid foundation for the course of therapy (Bordin, 1979; Connolly Gibbons et al., 2003; Gaston, 1990; Horvath & Greenberg, 1994; Horvath & Luborsky, 1993; Joyce, Ogrodniezuk, Piper, & McCallum, 2003; Snyder, Michael, & Cleavers, 1999; Strauss et al., 2006; Whisman, 1993). However, a strong early alliance influencing

outcome may depend upon different rating perspectives. Client ratings and to a lesser extent independent observer ratings of early alliance consistently predict outcome. Early therapist alliance ratings are less consistent predictors (Horvath & Symonds, 1991). Different rating perspectives are discussed further in the method section of this thesis.

Richmond (1992) and Wierzbki and Pekarik (1993) found that 40% to 50% of clients terminated therapy prematurely. Poor early alliance predicts client dropout (Blatt, Zuroff, Quinlan, & Pilkonis, 1996; Bordon, 1979; Botella et al., 2008; Constantino et al., 2002; Kokotovic & Tracey, 1990; Lingiardi, Filippucci, & Baiocco, 2005; Plotnicov, 1990; Tryon & Kane, 1995). Tryon and Kane (1995) found that clients who drop out of therapy give lower alliance ratings that those who stay until mutual termination. The initial phase of development of the alliance is thought to occur during the first five sessions, peaking most often by the third session (Eaton, Abeles, & Gutfreund, 1988; Horvath, 1993, 2001; Horvath & Greenberg, 1994; Horvath & Luborsky, 1993). Principe, Marci, Glick, & Ablon (2006) found that a client's return for a second session was significantly related to the Bond subscale of the WAI-S-C. A therapists application of techniques that convey trust, appreciation (Principe et al., 2006), warmth (Sexton, Littauer, Sexton, & Tommeras, 2005) and understanding (Tryon, 1990) tend to increase opportunities to improve alliance levels in an initial session. A strong bond may quell a client's anxiety and permit clients to participate in tasks or pursue goals that may stimulate anxiety (Pinsoff, 1994). Failure to engage with the therapist, inability to agree on tasks and goals of therapy or the lack of trust within the first three sessions may lead to disengagement from therapy (Horvath & Greenberg, 1994). Furthermore, if a collaborative relationship is not developed hope may turn into pessimism and if the therapist ignores the client's expectations rather than realistically renegotiating them the client may become anxious (Horvath & Greenberg, 1994). Positive alliance development may not take place immediately at the start of therapy, but Horvath and Greenberg (1994) maintained that the development of the alliance must be 'good enough' before therapeutic work can take place. This suggests that therapists need to pay attention to the alliance as soon as therapy begins (Castonguay et al., 2006). There is some evidence to suggest that strengthening the alliance during initial assessment is warranted (Ackerman, Hilsenroth, Baity, & Blagys, 2000; Busch et al., 2006; Hilsenroth & Cromer, 2007).

Once the workable levels of alliance have been developed, alliance patterns become more variable (Horvath & Greenberg, 1994). Relatively few studies have specifically

investigated the time course of the alliance. Bordon (1979) and Gelso and Carter (1985, 1994) proposed a U-shaped pattern to characterize the course of alliance in therapy. Horvath and Marx (1990) examined the alliance (as assessed by the WAI) at the session-by-session level in four clients and their two therapists. They found that therapists' combined ratings increased over the first 5 of 10 therapy sessions and then declined slightly, to recover marginally after the 7th session. Clients' ratings showed a linear increase over the therapy. Golden and Robbins (1990), in a similar study, reported a reverse pattern; the therapists alliance scores gradually increased over the 12 therapy sessions whereas his two clients' scores were lowest during the middle phase (5th-8th sessions) of therapy. In contrast, Kivlighan and Shaughnessy (1995) demonstrated that a direct linear pattern best accounted for client and therapist alliance scores over therapy. However, variability in methodologies of these studies may have contributed to the different alliance patterns. Horvath (1994a) argued that alliance measures, including the WAI may only be effective in measuring the alliance construct in the early stages of therapy. Once the tasks of a particular therapy start to impose unique demands and differential values on the dyadic relationship (Bordin, 1979), generic descriptors of the alliance measure may fail to accurately assess the alliance at the more differentiated levels (Horvath, 1994a).

'Third Variable' Confound in Alliance–Outcome Research

It is important to acknowledge that there may be a 'third variable' when attempting to understand the central features of alliance-outcome research (Crits-Christoph, Connolly Gibbons, & Hearon, 2006). There is a growing body of literature on factors that mediate or moderate the alliance in therapy including: client personal characteristics (Connolly Gibbons et al., 2003; Constantino et al., 2002; Hardy et al., 2001; Hersoung, 2001; Joyce et al., 2003), therapist personal characteristics (Ackerman & Hilsenroth, 2001, 2003; Barber et al., 2006; Hersoung, 2001; Luborsky et al., 1985), quality of object relations (Bordon, 1994, Kivilighan & Schmitz, 1992; Mallinckrodt, 1992; Piper et al., 1991), client intrapersonal resources, such as 'psychologically mindedness' and motivation (Roth & Fonagy, 1996), interpersonal histories and functioning of therapists and clients (Connolly Gibbons et al., 2003; Dinger, Strack, Leichsenring, & Schauenburg, 2007; Hillard, Henry & Strupp, 2000; Horvath, 1994a), interactional effects (Henry & Strupp, 1994; Talley, Strupp, & Morey, 1990), client expectations of session usefulness (Connolly Gibbons et al., 2003; Joyce & Piper, 1998), therapist experience and training (Hersoung et al., 2001; Kivlighan, Patton, & Foote, 1998; Mallinckrodt & Nelson, 1991) and the environment or

events in the individuals life (Luborsky et al., 1983; Needles & Abramson, 1980). Developing an alliance may also be confounded by health, legal, economic and social problems (Barber et al., 1999; Hersoung, et al., 2001).

There have been numerous investigations of the impact of severity of the client's disorder on the subsequent development of the alliance, but the results are mixed. While some studies have found the higher pretreatment symptom severity predicted lower alliance ratings (Eaton et al., 1998; Gaston et al., 1998; Hersoung et al., 2002; Zuroff et al., 2000), others have demonstrated that alliance ratings and pretreatment severity are not significantly related (Gaston et al., 1991; Luborsky et al., 1983; Joyce & Piper, 1998). Pretreatment depressive symptoms, and in particular negative cognitions and dysfunctional beliefs, have been associated with poor ratings of the therapeutic bond, but were unrelated to the goals and tasks components of the alliance (Rector et al., 1999). Clinical experience suggests that anxiety and depression comorbidity would influence the quality of the alliance (P. Merrick, personal communication, 11th September 2009), but research in this area is sparse. However, there is emerging literature surrounding comorbid personality disorders and alliance and drop out behaviour (Benjamin & Karpak, 2001; Bond, Banon, & Grenier, 1998; Lingiardi, Filippucci, & Baiocco, 2005; Muran et al., 2009; Muran, Segal, Samstag, & Crawford, 1994). Furthermore, there is evidence that a client's ability to engage in positive interpersonal relationships is one pre-treatment factor that may influence the development of a positive alliance (Marmar et al., 1989).

Summary

In summary, variability in findings regarding predictive significance of the alliance across different studies might reflect: properties of different diagnoses, varying inclusion and exclusion criteria, client suitability for therapy, varying diagnostic systems, the different demands of theoretical orientation of therapy, different treatment length, differential sensitivity of alliance measures, different outcome measures, session measured, how many sessions are measured, various sources of alliance ratings (i.e., client, therapist or observer rating), qualifications and training of raters, whether archived data is audiotaped or videotaped, varying data analyses, concurrent medication and in some cases, age. Methodological issues may also have added to the inconsistency of these findings and could include: small sample sizes, low statistical power and design sensitivity, 'third variables', lack operational definition of the concept itself within a particular theoretical orientation and confusion over the use of multiple alliance terms. The inconsistency

between the earlier findings and more recent findings may also be due to the natural evolution of more sophisticated statistical programmes and methods being developed for data analysis.

Treatment integrity is the degree to which an intervention is delivered as intended (Kazdin, 2003; Perepletchikova, Treat, & Kazdin, 2007; Waltz, Addis, Koerner, & Jacobson, 1993; Wampold, 1997). Effective treatment integrity, such as specifically training therapists for the study, supervision, and the utilization of experienced therapists and independent assessments of adherence and competence using randomly selected taped sessions, was more frequently considered and documented in the more recent alliance-outcome studies (Cuijpers et al., 2008; P. Merrick, personal communication, 11th September 2009; Valentine, 2005). While some of the alliance-outcome studies mentioned in this chapter utilized either audio or video taped sessions for supervision and/or assessed therapist adherence and competence within treatment procedures (Gaston et al., 1991; Gaston et al., 1998; Wilson et al., 1999; Klien et al., 2003; Strauss et al., 2006), only five of the 25 alliance-outcome studies examined between 1990 and 2006 utilized adherence and competence scales (Barber et al., 2006; Dunn et al., 2006; Feeley et al., 1999; Loeb et al., 2005; Stiles et al., 1998). Tang and DeRubeis (1999) utilized data from the Elkin (1989) study (i.e. NIMH TDCRP) and the Hollon et al. (1992) study. However, only the Hollon et al. (1992) study utilised an adherence and competence scale. Zuroff and Blatt (2006) also utilized data from the NIMH TDCRP. While it is documented that therapist competence was monitored by trainers/supervisors throughout treatment (Hill, O'Grady, & Elkin, 1992), it is unclear whether an adherence and competence scale was utilized throughout treatment in the Elkin (1989) study. Questions have been raised about the quality of CBT provided in the NIMH TDCRP. One of the three research sites in the study had poorer CBT outcomes and two of the CBT therapists were inexperienced and they received less frequent supervision (Hollon & Beck, 2004; Ilardi & Craighead, 1994). This can have serious implications for inferences drawn from the study between the relationship of the alliance and outcome (Kazdin, 2003; Perepletchikova, et al., 2007; Wampold, 1997). If therapist adherence and competence to treatment procedures are not measured and then documented there is uncertainty as to whether interventions have been implemented as described or whether the therapist has the capability to effectively implement the treatment (Nezu & Nezu, 2005; Perepletchikova, et al., 2007). Therefore, it is unclear whether the alliance process has been measured accordingly within the theoretical orientation(s) of the above mentioned alliance–outcome studies.

To date, researchers have mostly questioned the causal relationship between alliance and symptomatic change. Casual direction (if any) of the alliance-outcome relationship has not been clearly established. However, because of the complex nature of the relationship between alliance and symptomatic change it may now be more appropriate to investigate the sequential relationship between these two variables. Barber et al. (2000) reported that the alliance and outcome may be so intricately intertwined that they may amplify each other rapidly and may need to be measured every session. Notably all the abovementioned alliance-outcome studies employed between-group research designs and utilized either correlational or regression analyses. Therefore, this issue has not been explored on a individual level. However, different kinds of evidence are needed to contribute to the scientific process of developing useful and meaningful knowledge.

Recently, there has been much discussion surrounding frequency of measurement and statistical analyses of therapeutic processes. Some leading researchers in this field propose that simple pre-post outcome designs and pre-mid-post measures of change do not provide enough information or adequately capture the sequencing of therapeutic processes and outcome. Change may be gradual and linear or discontinuous and non linear and more frequent measurement allows better understanding of what facilitates and inhibits change. Some early case studies have utilized repeated measures to investigate the time course of alliance. However, to date, measuring the alliance and depression severity in every session to determine temporal relations (e.g., unidirectional, bidirectional, concomitant changes) within CBT has not been attempted. Statistical techniques such as growth curve modeling, growth mixture modeling and dynamical systems modeling are recommended to determine mechanisms underlying change (Laurenceau et al., 2007). Furthermore, research determining temporal relations between Goal, Task, and Bond components of the alliance and outcome is scant.

It is thought that the alliance forms early in treatment and is predictive of outcome. Recent findings, albeit relatively scant, suggest that the alliance may develop during initial assessment. Furthermore first impressions may count. However, it is unclear whether alliance measures including the WAI, can accurately assess the alliance in later sessions of CBT. While briefer versions of WAI have been developed more recently, the items of the WAI were developed over 25 years ago. Therapies such as CBT continue to evolve and as with other instruments items of the WAI may need up-dating (Tryon, Blackwell, & Hammel, 2008). Chapter 5 presents the aims and research questions of the current research.

Chapter 5: Aims and Research Questions of the Current Research

Overview

The following chapter details and describes the aims and the research questions of the current research.

Aims

The overall aim of the current research was to investigate the role of the alliance and symptomatic improvement within CBT for depression in order to contribute to enhancing the effectiveness of this therapy. The study of psychotherapy processes has been termed as a 'Rashomon' experience (Mintz et al., 1973) and there is a lack of consistency within methodologies and findings across alliance-outcome studies. Therefore, a second aim of the current research was to understand the alliance process and its measurement and its theoretical and empirical underpinnings in order to make careful methodological decisions. Essentially, it was important for the current research to address four key methodological considerations in order to investigate the role of the alliance and symptomatic improvement within CBT for depression.

The first methodological consideration of the current research was to review theoretical and empirical support for the alliance within psychotherapy in order to conceptualise the alliance within CBT and in particular align it with the relational quality of the collaborative empiricism (Beck et al., 1979). Theoretical and empirical support for the alliance within psychotherapy has often been reviewed, but within alliance-outcome research there has been little written report conceptualizing the alliance construct of the therapeutic modality utilized in the study. The alliance has been identified as a determinant of outcome within psychotherapy, regardless of theoretical orientation. However, while the alliance, in general may be similar across modalities, governing principles for defining alliance within each theoretical orientation varies and various specific aspects of the alliance may be emphasized, formulated or used differently. CBT emphasizes components of Goal, Task and Bond and appears to fit with Bordin's (1979) model of the alliance.

The second methodological consideration was to determine the most appropriate alliance scale to measure the alliance within CBT. Currently, as outlined in Chapter 3, there are at least eleven commonly used instruments available that vary in perspective (observer, therapist, or client). Each differs in its theoretical basis and conception of the relationship. Furthermore, most alliance–outcome research has assumed that the alliance is a one factor construct and has utilized the sum of the alliance scores for their analyses. However, there is a growing body of literature surrounding the possibility that the alliance may be composed of both a general higher order multifaceted construct and several more specific aspects or components (Andrusyna et al., 2001; Hatcher & Barends, 1996; Hatcher & Gillaspy, 2006; Tracey & Kokotovic, 1989). The current thesis reviewed the four most frequently used alliance scales and then the factor structure of the various versions of the WAI to determine the most appropriate instrument to measure the alliance and it's components within CBT.

The third methodological consideration was to determine a research design that would address the question of what is the intertwined and sequential relationship between the alliance process and symptomatic improvement. The current research reviewed alliance-outcome literature to determine the most appropriate research design to investigate temporal relations (e.g., unidirectional, bidirectional, concomitant changes) between the alliance and symptomatic change within CBT for depression. Prior work and theory was utilized to specify an appropriate measure rate and interval (Laurenceau et al., 2007). Most alliance–outcome studies have utilized either a pre-post or pre-mid-post research design to assess alliance and outcome. The alliance was either measured in the midst of treatment and correlated with symptom change from the beginning to the end of treatment (Castonguay et al., 1996; Gaston et al., 1998; Krupnick et al., 1996; Safran & Wallner, 1991) or it was averaged across the duration of the treatment and then this averaged measure was related to overall symptom change (Gaston et al., 1998; Krupnick et al., 1996). More recently, either alliance measures have been taken early in treatment, when therapeutic change presumably has not yet taken place and then related to outcome or evidence is provided by studies in which early change in symptoms is statistically controlled and measures of alliance continued to predict change (Barber et al., 2000; Klein et al., 2003). However, these methods may not adequately capture the slope or the direction of change of these variables. Now, it is recommended that process researchers increase the precision in which they study change by including more frequent assessments of symptoms and process mediators over the course of treatment (Laurenceau et al., 2007). Given the

findings it was decided that the current research would employ a single-case research design and multiple assessments to measure alliance and depression severity every session for the first ten sessions of therapy to gain greater understanding of the temporal relations between alliance and symptomatic change and the temporal relations between Goal, Task and Bond components of the alliance and symptomatic change.

The fourth methodological consideration was to train raters to rate the WAI-SR-O and establish inter-rater reliability for the current research. Inter-rater reliability establishes the extent of consensus on the use of an instrument by those who administer it (Shrout & Fleiss, 1979). Judgments made by humans are particularly prone to to measurement error and this can seriously affect interpretations of the findings (Shrout & Fleiss, 1979; Thorndike, 1920). In order to increase the inter-rater reliability of the current research, guidelines for rating the Working Alliance Inventory were sought from researchers overseas. Raue and his colleague's (1991) provided a set of guidelines to rate the original 36 item WAI. These guidelines were modified and expanded to rate the WAI-SR-O within CBT. The guidelines and their development are described in Chapter 6. A rater reliability study was conducted to provide a forum to train the raters for the current research and establish inter-rater reliability. Training the raters and the two stages of the rater reliability study are outlined in Chapter 7.

In summary, the current research addressed four key methodological considerations including: 1) conceptualise the alliance within CBT and align it with the relational quality of the collaborative empiricism, 2) determine the most appropriate alliance measure to assess the alliance and its components within CBT, 3) determine and utilize a research design that would address temporal relations between the alliance and symptomatic improvement and temporal relations between Goal, Task and Bond components of the alliance and symptomatic change and 4) train raters to rate the WAI-SR-O and establish inter-rater reliability. Addressing these methodological considerations was necessary to provide the platform to investigate the role of alliance and symptomatic change within CBT for depression.

Research Questions

The current research addressed four research questions in order to investigate the role of alliance and symptomatic change within CBT for depression. The first two research questions are replicated from previous alliance-outcome studies. However, the current

research utilized a different approach, that is a single-case research design and measurements of alliance and depression severity every session, to further understand what has already been investigated in the past and further inform the role of alliance and symptomatic change within CBT for depression. Research questions 3 and 4 evolved from Barber et al's. (2000) suggestion that alliance and symptomatic change may be so intricately intertwined that they may rapidly amplify each other. Previous alliance–outcome studies have mostly investigated whether alliance precedes symptomatic change or symptomatic change precedes alliance to seek causal change. However, because of the complex nature of the relationship between alliance and symptomatic change it may now be more appropriate to broaden these investigations to include investigations of the sequential relationship between these two variables (Barber et al., 2000).

1. Does a strong early alliance predict a positive outcome?

The current research investigated whether a strong early alliance was predictive of a positive outcome. Several investigations have supported early alliance being a strong prognosticator of outcome (e.g., e.g., Hartley & Strupp. 1983; Henry & Strupp, 1994; Hersoug et al., 2002; Horvath, 2000; Horvath & Luborsky, 1993; Loeb et al., 2006; Klien et al., 1993; Kokotovic & Tracey, 1990; Marzali, et al., 1999; Reandeau & Wampold, 1991; Strauss et al., 2006). Based upon current literature it was expected that a strong early alliance will be predictive of a positive outcome.

2. Does alliance precede symptomatic change?

Several early studies have reported that symptom improvement predicts a subsequent increase in the alliance (eg., DeRubeis & Feeley, 1990; Feeley et al., 1999; Safran, & Wallner, 1991). However, more recent research has suggested that alliance may precede symptom improvement (e.g., Busch et al., 2006; Hilsenroth & Cromer, 2007; Klien et al., 2003; Zuroff & Blatt, 2006). Most researchers conducting alliance-outcome studies have tended to view symptomatic improvement as a consequence of specific therapeutic intervention and sought causal direction. This thesis investigates the role of symptomatic change in the context of a change in depressive symptoms, regardless of whether this change has been caused by specific intervention. However, for this particular research question the current research will also consider examining symptomatic change in the context of symptomatic change being a consequence of specific therapeutic intervention and interpreting its findings accordingly. Based on more recent research it was expected that alliance will precede symptomatic change.

3. What is the intertwined and sequential relationship between alliance and symptomatic change?

The current research questioned the intertwined and sequential relationship between alliance and symptomatic change. Partialing out prior change in depression, Barber et al. (2000) demonstrated that the alliance predicted subsequent change in depressive symptoms, but greater symptomatic improvement through session 5-10 led to higher levels of alliance, and that, in turn, alliance was associated with further symptomatic improvement. These authors suggested that that there may be reciprocal effects between change in the alliance and change in depressive symptoms which needed further investigation. Therefore, based on these findings it was expected by the current research that alliance plays a role in the change of depressive symptoms and that symptomatic change influences the development of the alliance.

4. What is the intertwined and sequential relationship between Goal, Task and Bond components of the alliance and symptomatic change?

The current research questioned the intertwined and sequential relationship between Goal, Task and Bond components of the alliance and symptomatic change. There has been a dearth of research on the relationship between the various components of the alliance (e.g., Goal, Task and Bond) and their relationship/s with outcome. Rector et al. (1999) found that perceived agreement between the client and therapist on the goals and tasks of therapy predicted change in dysfunctional beliefs in CBT, but that a strong therapeutic bond appeared to provide the context in which changes in dysfunctional beliefs resulted in significant symptom change. Although the bidirectional relationships between the alliance components and cognitive change could not be determined in that study, it was expected in the current research that agreement on the tasks and goals would be important in facilitating change, whereas symptom change would be greater in the context of a strong bond between the client and therapist.

Chapter 6 details the methodology of the current research.

Chapter 6: Methodology - Current Research

Overview

The current research utilized archived session data recorded on digital video discs (DVDs) from the Depression Outcome Study conducted at the School of Psychology at Massey University, Albany, Auckland. The participants of the Depression Outcome Study completed the BDI-II at the beginning of every therapy session and this data was used in the current research to assess symptomatic change. Chapter 6 is divided into two sections. The first section pertains to the Depression Outcome Study and describes the characteristics of the study's first ten participants, whose session data and BDI-II scores were utilized by the current research. It also outlines selection criteria, initial telephone assessment, pre-treatment measure, therapist training and supervision, therapist adherence and competency measures, client suitability measure and treatment received by the ten participants. The Depression Outcome Study utilized The Attributional Style Questionnaire and the Social and Occupational Functioning Assessment Scale as outcome measures, but the current research did not utilize the scores of these measures and therefore they will not be described in this thesis. However, this thesis will describe the psychometric properties of the BDI-II. The second section of this chapter pertains to the current research and describes the process measure. It outlines the development of the WAI-SR-O and the guidelines to rate the WAI-SR-O and discusses expert feedback pertaining to the development of these guidelines. It also details the characteristics of the raters and the rating of the WAI-SR-O. This chapter concludes with ethical considerations and discusses the single-case research design and data analytic procedures relevant to the current research.

Participant Characteristics

The sample was drawn from the bigger Depression Outcome Study which was investigating a number of process and outcome variables in psychotherapy. Participants for the Depression Outcome Study were recruited by placing articles and advertisements in the local newspaper and phoning and mailing brochures to mental health providers and general practitioners in the Auckland area. The sample of participants for the current research was made up of the first ten participants, ages 18 to 65 years. Nine of the participants were

Caucasian and one was of Caucasian and Polynesian descent. They (8 female and 2 male) met selection criteria of having a major depressive episode (DSM-IV-TR; 2000) for the first time, were proficient in reading, writing and conversing in English, free from taking CNS acting drugs (except the occasional hypnotic or the oral contraceptive), did not meet diagnostic criteria for substance abuse, psychotis, borderline personality disorder, and were able to be safely managed with outpatient psychotherapy. They were not receiving any other form of psychotherapy. Demographic characteristics of each participant are summarized in Table 6.1.

Table 6.1: Demographic Characteristics of the Participants.

Client	Gender	Age	Depression Severity at Pretreatment Assessment	Comorbid Axis I Disorder assessed by CIDI
1	Female	25 years	Moderate	PD without agoraphobia.
2	Male	62 years	Mild	
3	Female	38 years	Severe	BN (from 11 yrs old), PTSD (from 8 yrs old), GAD, hypochondrias.
4	Female	46 years	Severe	PTSD, GAD
5	Female	37 years	Mild	
6	Female	51 years	Moderate	GAD
7	Female	31 years	Moderate	PD without agoraphobia, GAD
8	Female	47 years	Severe	OCD, PTSD, GAD
9	Male	57 years	Severe	PD without agoraphobia, GAD, PTSD
10	Female	49 years	Severe	GAD

Note: PTSD‑post traumatic stress disorder, GAD‑generalized anxiety disorder, OCD-obsessive-compulsion disorder, BN–bulimia nervosa, PD–panic disorder.

No association has been found between the alliance and client demographics (Gaston, Marmar, Thompson, & Gallagher, 1988; Marmar et al., 1989), with the exception of one association between the alliance and level of education (Marmar et al., 1989). When reviewing literature pertaining to the influence of client demographic variables on

psychotherapy it appeared age and gender were not important in either therapy retention or treatment outcome (MacDonald, 1994; Petry, Tennen, & Affleck, 2000; Sledge, Moras, Hartley, & Levine, 1990). Race related client and therapist variables (e.g., race matching between therapist and client) did not predict therapy outcome (Jones, 1978; Jones & Zoppel, 1982). Depression severity and comorbidity of the ten participants will be further commented on in the results section. Throughout the remainder of this thesis the participants of the Depression Outcome Study will also be referred to as clients.

Procedure

Initial Assessment

Referrals for the Depression Outcome Study were initially screened over the telephone by a doctoral-level clinical psychology student (NF) who was also experienced in telephone interviewing and specifically trained in assessing the study's inclusion criteria. Prospective participants were given full written and verbal information about the study and asked to provide written consent (Appendix A). Participants then underwent a detailed pretreatment assessment administered by a study coordinator (two PhD-level psychology students were trained to administer pretreatment assessments - MM, ME) before being scheduled to meet with the next available therapist for a second clinical assessment session.

Pretreatment Assessment

Pretreatment assessment included a structured clinical interview to determine the presence or absence of DSM-IV Axis I disorders (Composite International Diagnostic Interview; CIDI) and BDI-II (A. T. Beck et al., 1996) to assess depression severity. Depression severity was further assessed at the beginning of every session (i.e., twice weekly during approximately the first half of the treatment and once-weekly during the second half) to not only assess symptom change, but to also ensure the ongoing safety of the participants within the study. The Depression Outcome Study also utilized The Attributional Style Questionnaire (ASQ; Peterson et al., 1982) to assess current attributions, and the Social and Occupational Functioning Assessment Scale (SOFAS; Hay, Katsikitis, Begg, Da Costa, & Blumenfield, 2003; Hilsenroth et al., 2000 to evaluate current functioning at pretreatment and sessions 5 and 8. The current research has followed the methodology of previous alliance process-outcome studies and only utilized archived BDI-II scores to measure symptom change. Each participant was further assessed by their therapist at a second initial clinical assessment session to ensure that the participant meet inclusion criteria and to rate their treatment suitability with the Suitability for Short Term Cognitive

Therapy scale (SSCT: Safran & Segal, 1990). Each participant was interviewed by their therapist approximately a week after the pretreatment assessment and if they meet inclusion criteria the first therapy session was usually scheduled for the following week. There was no waiting list.

Therapists

CBT (Beck et al., 1979) for depression was conducted by three female doctoral level clinical psychology students (CO, JD, RF) who were also researchers in the Depression Outcome Study project. The three therapists were Caucasian and their ages ranged between 29 and 50 years old. The current researcher was one of the therapists of the first ten participants in the study.

Therapist Training

Therapists completed training requirements before providing therapy. This training process included a two day intensive training workshop on CBT for depression by Professor Keith Dobson (University of Calgary) and two five day intensive postgraduate training courses for Cognitive Behaviour Therapy at Massey University (e.g., 175.761 – Theory and Practice of Cognitive Behaviour Therapy and 175.762 – Cognitive Behaviour Therapy for Depression). An additional five training workshops focused on the use of homework assignments in CBT and learning the guiding model for reviewing, designing and assigning homework (Kazantzis et al., 2005) was conducted by the primary investigator (NK) of the Depression Outcome Study. Taped practice sessions in which the therapists took turns to role-play either the therapist or client were reviewed and critiqued. All therapists attained 100% adherence on the Homework Adherence and Competence Scale (HAACS; Kazantzis, Wedge, & Dobson, 2004) before being allowed to provide therapy. They also demonstrated competence in delivering CBT for depression as defined by their clinical supervisor's ratings on the Cognitive Therapy Scale (CTS; Young & Beck, 1980), the international standard assessment scale for assessing competence in cognitive therapy. A total score of 40 or greater on the CTS represents the standard threshold of acceptable competence in Cognitive Therapy delivery (Dimidjian, et al., 2006; Elkin, 1999). The therapists completed a one day workshop on initial assessment and risk assessment training.

Therapist Supervision

Clinical supervision was conducted by a formally trained independent clinical psychologist with 11 years clinical and CBT experience (RV). Therapists received ongoing assessment

in their competence for delivering CBT for depression from their clinical supervisor utilizing ratings on the CTS (Young & Beck, 1980). There ratings provided standards for quality control monitoring or therapist and client safety and treatment integrity. Therapist supervision was structured into two phases, training phase and clinical phase. During the training phase, that is, up to session 8 or first 4 weeks of therapy with the first set of three clients per therapist, three CTS ratings for each therapist were rated by the clinical supervisor. During the clinical phase five CTS ratings were scheduled. In total, each therapist was to be rated on 8 CTS ratings. The sample participants for the current research were the first set of three clients per therapist, that is in the training phase of the therapists, plus one participant from the clinical stage of one of the therapists. This last participant terminated therapy after six sessions. Regular group supervision comprised of presenting a session DVD snippet and conceptualization of the client, discussion surrounding treatment plan and feedback from the clinical supervisor. All therapists gained above the score of 40 for the three CTS ratings needed in the training phase. Homework protocol adherence ratings were conducted by the two study coordinators and then monitored by the clinical supervisor. Therapists received ongoing assessment of their adherence in the use of homework assignments in CBT for depression to ensure treatment integrity.

Treatment

Participants received up to 20 sessions of CBT for depression (A. T. Beck et al., 1979), utilizing a guiding model to facilitate reviewing, designing and assigning homework over a 16 week period. Consistent with prior research on CBT for depression, the first 8 sessions were offered twice weekly (A. T. Beck et al., 1979; J. S. Beck, 1995). Follow-up booster sessions occurred at 2 months and 6 months post-treatment. Therapy was free of charge. All therapy sessions were DVD-recorded and then archived.

Measures

Pretreatment Assessment Measure

Composite International Diagnostic Interview (CIDI)

The Composite International Diagnostic Interview or CIDI , is a fully structured, highly standardized interview that maps the symptoms elicited during the interview onto DSM-IV and ICD-10 diagnostic criteria, and reports whether the criteria is satisfied (Andrews & Peters, 1998). The instrument contains 276 symptom questions to evaluate symptom severity, as well as questions assessing help-seeking behaviour, psychosocial impairments and other episode-related questions (Wittchen, 1994). The CIDI is available in lifetime and

12-month versions, and in paper-and-pen and computerized forms. In a review of CIDI Version 1 field trials Wittchen (1994) reported test-retest agreement of 86%, test-retest kappa values of .66, inter-rater agreement of 100%, and inter-rater kappa values of .97 for the DSM-III-R diagnosis of major depression, single episode. Studies involved two independent clinical and non-clinical interviewers. The Composite International Diagnostic Interview (CIDI) Version 2.1 was developed by the World Health Organisation (1997) and the National Institute of Health (Andrews & Peters, 1998). Increasingly, studies are employing structured interviews to assess the prevalence of psychiatric illness (Komiti et al., 2001). Computerised versions of structured interviews have been reported as improving standardization of diagnosis, eliminating clinician bias and also offering high reliability and consistency of administration. They are cost effective and time efficient, and eliminate errors in data entry and scoring (Erdman, Klien, & Blouin, (1985). However, studies (Komiti et al., 2001; Rosenman, Korten, & Levings, 1997) suggest that agreement between the computerized version of the CIDI (CIDI-Auto) and experienced clinicians is poor for major depression. Caution is required when viewing results from the CIDI-Auto (Komiti et al., 2001).

Outcome Measure

Beck Depression Inventory –II (BDI-II)

A screening device, rather than a diagnostic tool, The BDI-II is a widely used 21-item patient-report measure of depression severity with possible scores from 0 to 63 (Beck, et al., 1996). Each item represents a symptom characteristic of depression, such as sadness, guilt, suicidal thoughts, and loss of interest (Beck et al., 1996). With regard to internal consistency, coefficient alpha estimates were found to be .92 and .93 for a psychiatric outpatient sample and .93 for a sample of college students. Test-retest reliability was estimated based on 26 outpatients who completed the BDI-II twice, one week apart. This stability correlation was .93 (Beck et al., 1996; Dozois, Dobson & Ahnberg, 1998). The BDI-II significantly correlates with an earlier version of the inventory, the BDI-I (Dozois et al., 1998). With regard to concurrent validity the BDI-II was found to correlate .71 with the Hamilton Rating Scale for Depression (Hamilton, 1960) (Beck et al., 1996). Nineteen items include a 4-point scale ranging from 0-3, representing ascending levels of severity, form the absence of a given symptom (e.g., "I do not feel sad") to an intense level (e.g., " I feel so sad or unhappy that I can't stand it"). The remaining two items (i.e., changes in sleep patterns, changes in appetite) allow the respondent to indicate increases or decreases in these behaviours (Beck et al., 1996).

There are no agreed methods on how to establish clinical change on the BDI. There has been two decades of discussion and there is still no resolution (Beck et al., 1996). Furthermore, most methods require local normative data. Studies providing norms and/or psychometric properties for the BDI in New Zealand populations are few (eg., Knight, 1984; Siegart, Walkey, & Taylor 1992). Knight (1984) provided age and sex-specific norms for the 13-item short form of the BDI based on data collected in a general health survey in the Borough of Milton, a community 54Km south of Dunedin. With an internal consistency reliability of .81, they found that females scored more highly than males. However, these studies would have utilised the earlier BDI Scale. For the BDI there are agreed upon conventions for establishing what is clinically meaningful in terms of BDI change. For research purposes, patients are usually classified as non-depressed if their BDI scores fall below the clinical cutoff of 13, "minimal depression" (Beck et al., 1996). Scores range from 14–19 for mild depression; 20–28 for moderate depression and 29–63 for severe depression (Beck et al., 1996).

Therapist Adherence and Competency Measures

Cognitive Therapy Rating Scale

It was critical to consider the extent to which the therapists were engaged in Beckian CBT and the Depression Outcome Study utilized The Cognitive Therapy Scale (CTS; Young & Beck, 1980) to assess the competence with which CBT was delivered. However, rater's skill can be a confounding variable and some have stricter criteria than others. Novice raters can be more rigid and mechanical than more experienced raters who may have a more global and integrated picture (P. Merrick, personal communication, 11th September 2009). The therapists of the Depression outcome Study were rated by only one rater, so the quality of ratings were consistent and the rater had many years of clinical and supervisory experience within CBT. Currently the CTS is the most widely used standardized instrument to assess the competence within CBT.

The CTS is an 11-item instrument designed to measure the quality of treatment delivery for cognitive therapy over two domains; General Therapy Skills (e.g., agenda setting, interpersonal effectiveness, collaboration, pacing, use of feedback and understanding); Conceptualisation, Strategy and Technique (e.g., focusing on key cognitions and behaviours, homework, strategy for change, and guided discovery). Each item is rated using a 7–point Likert scale ranging from 0-6. It demonstrates reliability when used by expert raters (Dobson, Shaw & Vallis, 1985; Vallis, Shaw, & Dobson, 1986) and is

sensitive to variability in the quality with which a cognitive therapy protocol is administered (Vallis et al. 1986). Inter-rater reliability, as assessed by Pearson product moment correlations, for the CTS individual items ranged from 0.54 to 0.87, with an average across 21 therapy sessions of 0.69. The inter-rater correlation of the total score was 0.94 (Dobson et al., 1985).

<u>Homework Adherence and Competence Scale (HAACS)</u>

The Homework Adherence and Competence Scale (HAACS; Kazantzis et al., 2005) was developed to determine the extent of therapist adherence and competence in the design, assignment, and review of homework as described in the homework protocol. The HAACS assesses CBT therapists' in-session behaviours in using homework assignments, including therapists' activities in designing, assigning, and reviewing homework (Kazantzis et al., 2005). An initial study examined data from two raters using a combined adherence and competence scale (16 items) to rate a brief 8-session protocol of CBT for anxious and depressed clients. Item analysis, revision, and deletion were conducted and resulted in a 19-item revised HAACS scale with separate ratings for adherence and competence (Wedge, 2005). A second study examined data from five raters using the revised HAACS to rate four DVD-recorded therapy sessions from a National Institute of Mental Health study of cognitive therapy for depression. The two pilot evaluations in second study resulted in further refinement to HAACS items (Wedge, 2005). A further study has examined data from two doctoral-level raters using the HAACS to rate 76 DVD-recorded therapy sessions from the clinical phase of the NIMH trial. Results indicate that the HAACS has strong levels of inter-rater reliability for adherence (.77) and competence (.81) and criterion validity (Munro, 2006). Inter-rater reliability for the 28 clients in the Depression Outcome Study was 0.93.

Client Suitability Measure

<u>Suitability for Short Term Cognitive Therapy</u>

The Suitability for Short Term Cognitive Therapy (Safran & Segal, 1990) is a therapist-rated ten-item questionnaire, providing an overall and item by item measure of patient suitability for cognitive therapy (Safran, Segal, Vallis, Shaw, & Samstag, 1993). The items are rated on a 1-5, with half point ratings, point scale. These scales are anchored so that 5 indicates the best prognosis and 1 the worst. The items focus on the patient's ability to access thoughts; differentiate emotions; accept personal responsibility for change; identify and work with cognitive rationale; display alliance potential (in-session evidence); display

out-of-session alliance potential (i.e., maintain good relations with family, friends and peers); be focused; display optimism regarding positive outcome; and avoid defensive and avoidant strategies. The tenth item relates to problem chronicity of personal responsibility for change, compatibility with cognitive rationale, and alliance potential. Dunn et al. (2006) found that lower client suitability as measured by the Suitability for Short Term Cognitive Therapy Scale (Safran & Segal, 1990) at session three predicted lower alliance.

Process Measure – Current Research

Working Alliance Inventory – Shortened Revised (WAI-SR)

The WAI–SR (Hatcher & Gillaspy, 2006) is a revised version of the WAI-S (Tracey & Kokotovic, 1989). These scales and their psychometric properties were described earlier in Chapter 3. Permission was gained from A. O. Horvath to use the WAI-SR in the current research (A. O. Horvath, personal communication, 21st November, 2006).

Development of the WAI-SR-O Scale

The WAI-SR-C was modified to develop the WAI-SR-O (see Appendix B) for the current research to analyse archived session data from The Depression Outcome Study. Items from the client version of the WAI-SR were replaced with corresponding items from the original observer 36-item version of the WAI. For example, item 9 of the WAI-SR-C, "*I believe _____ likes me*" was replaced with item 30 from the original observer 36-item version of the WAI, "*There is mutual liking between the client and the therapist*". The wording of these items was checked by another doctoral clinical psychology student (RF). Hatcher and Gillaspy's (2006) suggestion of a 5-point scale (i.e., 1 = seldom, 2 = sometimes, 3 = fairly often, 4 = very often, 5 = always) was incorporated.

Rating the Working Alliance Inventory

Although all three perspectives (i.e., client and therapist self reports, objective observer) are somewhat predictive of outcome (Luborsky, 1994), there are between-perspective differences with regard to both the judgment of the quality of alliance and the link with therapy outcome (Bachelor, 1991; Barber et al., 1999; Ceceroc et al., 2001; Hatcher, Barends, Hansell, & Gutfreund, 1995; Horvath & Symonds, 1991; Mallinckrodt & Nelson, 1991; Ogrodniczuk et al., 2000; Tichoner & Hill, 1989). Horvath and Symonds (1991) (meta-analysis of 24 studies) showed client (mean effect size = .27 and observer (mean effect size = .23) evaluations were stronger than therapist judgment (mean effect size = -.03). These effect size values reflected the averages across all three rating sources of

outcome. With regard to the outcome ratings, it was revealed that client-rated outcome (r = .21) was moderately better predicted than therapist- (r = .17) or observer-rated (r = .10) outcome. Strong correlations were reported between client-rated alliance and client-rated outcome (r = .31) and observer-rated alliance and client rated outcome (r = .29). However, client–rated alliance and outcome may have been biased by inflated and converging self-assessments (i.e., halo effect). Therapist-rated scores were the least predictive of outcome (r = .22) (Horvath & Symonds, 1991).

Horvath (1994a) offered several potential explanations regarding therapist ratings being poor predictors of outcome which included: the alliance not being experienced in the same manner by client and therapist, therapists having to infer the beliefs or feelings of the client and finally therapists assessment of the alliance is not as "pure" as that of the client and may be influenced by a more global and theory-driven perspective of the treatment process and is more optimistic about the expectations of change. Theoretical knowledge of how the alliance should be developed or expectations of success in developing the alliance may alter the therapist's judgment of the relationship. Horvath (1994a) also argued that the therapist's investment in viewing themselves as empathic persons successfully engaged with clients may facilitate misconceptions and distortions of the relationship based on the therapists own relational dispositions.

Research has suggested that it may be more productive to focus on the client's subjective experience of the therapeutic encounter (Elliot et al., 1990; Lietaer, 1990, Orlinsky & Howard, 1986b). Emotions inform people about their environment's impact on them, trigger their action tendencies in response to that impact and provide a means to communicate with others (Watson & Greenberg, 1994). When rating a client version of an alliance scale clients rate the subjectively felt impact of being in a relationship with the therapist (Watson & Greenberg, 1994). Clients monitor their alliance with the therapist through a prereflective attention to their feelings and their sense of themselves in the session (Glaser & Strauss, 1967), which provides them with a sense of themselves in the relationship (Greenberg & Safran, 1984, 1987; Orlinsky & Howard, 1986b). However, when assessments are subjective it is possible that alliance-outcome relationships may be influenced by a "halo effect" where the alliance and outcome are rated by the same source. (i.e., if the client believes he/she has improved, it is more likely that he/she will rate the alliance more positively) (Fenton et al., 2001).

Observer versions of alliance instruments are not usually recommended because raters must make inferences about clients' thoughts, feelings attitudes and motives for which evidence would be hard to find (Elvins & Green, 2008; R. L. Hatcher, personal communication, 19th April, 2007; Hatcher & Barends, 2006; Hill & Lambert, 2004; Horvath & Greenberg, 1994). Furthermore, this complex inferential process may be biased by the rater's theory (Horvath & Greenberg, 1994). When rating bond components of the alliance such as empathy, raters may need to take into account the therapist's and client's tone of voice, openness of their body language, open expressions of rapport and concern, and appropriate displays of humour (de Roten et al., 1999; Raue et al., 1991). Goal and Task items are often based upon the assumption that 'agreement' is happening between the therapist and client (L. R. Fenton, personal communication, 1st September, 2007). For example, the rater observes whether the client and the therapist appear willingly and enthusiastically engaged in the therapeutic process of working on mutually agreed agenda items and that the client and the therapist appear willingly and enthusiastically in discussion about what will help the client accomplish their goals. Raters need to agree upon a group definition about how to define the item (L. R. Fenton, personal communication, 1st September, 2007).

Observers' data (i.e., videotapes or audiotapes) permits replication and yields information that is more objective (Horvath & Greenberg, 1994). Videotapes also allow nonverbal behaviours to be noted. The use of multiple raters (e.g., three raters) increases rater reliability (Shrout & Fleiss, 1979), protects against raters drift (Kazdin, 1977; Tsui, 1983), and counters the possibility of rater dropout if the study is conducted over an extended period of time (Gaston & Marmar, 1994). The use of a recalibration training session after the completion of 10 consecutive ratings also offers protection against raters drift (Gaston et al., 1998).

Research has demonstrated that inter-rater reliability estimates of the WAI-O range from .62 to .92 (Fenton et al., 2001; Hanson et al., 2002; Tichenor & Hill, 1989), which meet professional standards of acceptability (Cicchetti, 1994; Cicchetti & Sparrow, 1981; Fleiss, 1981).

Training for Independent Rating of the WAI-O

Client and therapist versions of the WAI are reported as being easy to administer; A.O. Horvath no longer maintains a "users manual" (A. O. Horvath, personal communication,

21st November, 2006). For more information on use and interpretation of the measure he directs researchers to his chapter; Horvath, A. O. (1994a). Empirical validation of Bordin's pantheoretical model of the alliance: The Working Alliance Inventory Perspective. In A. O. Horvath & L. S. Greenberg (Eds.), The Working Alliance: Theory, Research and Practice (pp. 109-128). New York: John Wiley & Sons.

Some researchers utilizing the WAI-O have found the instrument to be economical and that it does not require any rater training (Tichenor and Hill, 1989). However, others (e.g., Cecero et al. 2001; Raue et al., 1991) have pointed out the need for a thorough manual to be developed to guide raters to use the instrument. They found that extensive training was used to achieve adequate reliability on the WAI-O (Andrusyna et al., 2001; Cecero et al. 2001; Fenton et al., 2001). This training included going over each item as a group, figuring out which subscale it may belong to, discussing it and then rating practice tapes not associated with the actual study session. This group rating was lengthy, because often, there were so many different opinions and perspectives on items. Discrepancies were discussed as they tried to come to some consensus (L. R. Fenton, personal communication, 1st September, 2007; P. J. Raue, personal communication, 29th January 2008.). Ongoing reliability assessment (e.g., randomly selecting a session and each rater rating that session independently) and regular meetings to discuss ratings and discrepancies were established (P. J. Raue, personal communication, 29th January 2008.).

Development of the Guidelines for Rating the WAI-SR-O

It is difficult to operationalise the alliance, particularly client and therapist feelings and perceptions as opposed to their behaviours. Therefore a more concrete way for observers to rate this process was needed to increase inter-rater reliability. Initially, experts from overseas (e.g., R. J. DeRubeis, L. R. Fenton, L. Gelford, P. J. Raue and D. Strunk) were consulted about training packages they had used in their own studies for rating the WAI. Raue and his colleagues (1991) provided a set of guidelines that they had developed for the WAI-O to measure alliance in their CBT versus psychodynamic-interpersonal study. These guidelines were based on Bordin's model of the alliance (i.e., explanations about the Goal, Task, and Bond subscales) (P. J. Raue, personal communication, 9th May 2007). General guidelines explained that ratings were based on frequencies of occurrence within the session, but that intensity should also be taken into account, items referred to the beliefs and experiences of the client and therapist, and coding was based on both subjective and objective scores of the evaluation. Raters were asked to give the relationship the "benefit

of the doubt", so when rating positively-valenced items number 7 ("always") was taken as the starting point and scores were decreased as detractor were identified. Similarly when rating all negatively-valenced items, number 1 ("never") was taken as the starting point and scores are increased as negative signs are identified (Raue et al., 1991). Furthermore, raters were asked to rate the alliance, not the variables potentially responsible for it (e.g., the pathology of the client and/or the competence of the therapist). Guidelines for scoring specific items were also defined. For example, the item "*There is mutual liking between the client and the therapist*", was defined as "Mutual liking" involves having pleasant and warm interactions, being engaged, being sensitive, and showing personal interest in each other (e.g., the therapist should enhance or reinforce client expressions of personal values or outside interests, and the therapist should not forget important details about the client's life) (Raue et al., 1991).

The current research modified and extended Raue and his associate's (1991) guidelines to cover items of the WAI-SR-O and to facilitate the rating of the alliance within CBT. The newly developed instructions on defining Goal, Task and Bond items and rating guidelines were discussed thoroughly during group supervision with other team members (i.e., supervisors, doctoral and PhD clinical psychology students) of the Depression Outcome Study. For example, it was suggested to put the ratings and specific guidelines on one page, concepts such as 'mutual liking' and 'respect' and 'collaboration' and 'agreement' were clearly defined, and frequency and intensity ratings were developed more fully. Drafts of the guidelines were sent to international experts and their feedback was incorporated into the rating guidelines used for training the raters for the current research. Feedback from the participants from the training session (see Chapter 7 for more details of this feedback) was also incorporated into the final draft.

Expert Feedback

The experts did raise a few concerns about the guidelines and these were taken into consideration in the final draft. For example, the modifications of the guidelines had been specifically adapted to measure alliance within CBT and therefore they may be less useful for other therapies that do not utilize specific techniques such as setting an agenda or goals for a session (L.R. Fenton, personal communication, 2008). As established in Chapter Two the governing principles for defining alliance within each theoretical orientation varies and one of the objectives of the current research was to specifically capture alliance processes within CBT. While the WAI was a transtheoretical instrument and it has been used to compare the quality of the alliance in sessions of psychodynamic-interpersonal therapy and

CBT (Raue, 1995; Raue et al., 1993; Raue et al., 1997), Bordin (1979) did posit that different theoretical orientations would place different demands on to the alliance. Furthermore, the question had been raised whether brief general alliance scales would detect the collaborative processes between the therapist and the client within CBT (Hatcher, 2006). Therefore, it was deemed important to train the raters with specific guidelines that that aligned with the relationship quality of collaborative empiricism (Beck et al., 1979) to facilitate rating the alliance within CBT.

The experts also pointed out that there have been previous objections about counting methods for psychotherapeutic processes in the past and the danger of interpreting the frequency of instances too concretely (L.R. Fenton, personal communication, 1st September, 2007; L. Gelfand, personal communication, 4th February 2008; P. J. Raue, personal communication, 29th January 2008). The guidelines were also altered to provide levels of engagement to determine intensity of the alliance in Goal and Task items. Defining intensity in the Bond items was more subjective and the language of these items was changed to reflect the feelings and experiences of the client and therapist that may not necessarily be expressed verbally. The raters were also asked to consider whether any ruptures or strains in the relationship were investigated and repaired.

Finally, while as much guidance as possible is needed to ensure reliable ratings there is tension between being very specific versus the rater's overall clinical judgment (P.J. Raue, personal communication, 29th January 2008). Bernardin, Alvares, and Cranny (1976) found that scales with clarification statements at the anchor points had significantly less leniency errors and greater rater discriminability than scales without such statements. However, this study involved scale development of college teaching performance so their findings may not generalize to process research. The Cognitive Therapy Scale (CTS; Young & Beck, 1980) was developed to examine the relationship of therapist and patient variables to changes in competence within CBT. A revised version, the CTS-R (Blackburn et al., 2001) was developed to improve on the CTS by eliminating overlap between items, improving the scaling system and defining items more clearly. While more research is needed to be able make more reliable judgments about its proficiency (Keen & Freeston, 2008), the CTS-R is being used regularly to evaluate CBT trainees in the Newcastle Cognitive and Behavioral Therapies Centre and other Cognitive and Behavioral Therapies Centres throughout the United Kingdom (P. Armstrong, personal communication, 15th October, 2009). It is more detailed than the CTS and on one hand raters have found it more helpful

to have the more accurate guide, but on the other hand they found there was more room for discrepancy. It was also reported that two raters were often needed to rate a session and if these raters couldn't agree, a third rater also rated the session, which of course took more time. However, they have noticed an improvement in rating across the CBT training centres (P. Armstrong, 15th October, 2009). Therefore, having specific anchor points to guide raters rating the alliance may be beneficial in the current research.

To date, training people so that they can be more reliable raters of the alliance has not been addressed. Therefore, the current research project (see Chapter 7) investigated whether giving more information (i.e., guidelines developed for rating the WAI-SR-O) would lead to less variance of ratings and higher inter-rater reliability.

Final Draft of the Rating Guidelines (Appendix C)

Similar to the original guidelines (Raue et al., 1991) the newly developed guidelines described the Goal, Task and Bond dimensions of the alliance based on Bordin's model.

Goals were defined as the objectives of the therapist and client that specifically targeted change. Goals could include reduction in symptoms, improvement in interpersonal skills or relationships, awareness of intrapersonal conflicts, and development of new ways of thinking or behaving. Agreement was assessed according to the extent to which both client and therapist see the goals as important, clear and capable of being accomplished. The understanding of the benefits resulting from the change and degree of confidence in change could also be assessed.

Tasks were defined as therapeutic techniques or processes that help clients to increase awareness of their own thoughts, feelings, values and needs. Each task has a particular therapeutic goal. Examples of tasks are: support/reassurance, reflection, reformulation, cognitive restructuring, any type of reality testing, therapist suggestion/advice, information giving, homework assignments and role playing. Ratings on tasks are made according to how responsive the client is to task, and how responsive the therapist is to the client's need in suggesting the task. Therefore, the tasks should be seen as important, appropriate, and clear by therapist and client for high ratings. Agreement in these cases would depend on the degree of rejection (i.e., presence of thoughtful consideration of the technique) as well as their frequency.

Bond was defined as the mutual liking, respect, and trust between the client and the therapist. It was also characterized by therapist genuineness, openness, warmth, and

understanding. Client confidence in the therapist, client comfort, and reciprocal respect were also included. The bond was assessed through the client's tone of voice, amount and quality (i.e., degree of comfort) of client talk concerning intimate issues. Bond was also assessed through the therapist degree of comfort, non-defensiveness, accurate empathy (especially empathy validated by the client), and the mutual value placed on each others contributions.

A note was added that collaboration (i.e., both the therapist and the client participate in decision making, formulating ideas) in therapy requires agreement, whereas agreement (i.e., either the client or the therapist suggests an idea/task/goal and the other agrees) is possible without collaboration.

Rating guidelines explained and gave examples of how to rate frequency and intensity of alliance occurrence within a session. The guidelines also stated that all the items referred to the therapist and the client dyad, but emphasis was placed on the client's experience, and that rating the alliance involved both subjective and objective evaluation criteria. Ratings also were based on what is going on between the therapist and the client to determine whether they were working well together, had a good rapport, reciprocal relationship and if the client was working and getting value out of the sessions. Ratings were not based on the therapist's style of relating. Raters were asked to consider all relevant information from the entire session, but not to consider the pathology or other characteristics of the client when making ratings nor rate the technical skill or the competence of the therapist exhibited in the session.

Ratings on this form assumed that the relationship quality is present. Ratings were subtracted when a therapist alliance behavior would have been helpful, when a therapist alliance behavior was clearly unhelpful, or when a therapist alliance behavior was absent. Simultaneously, the intensity or strength of the affective bond and the ability of the therapist and the client to work together in therapy needed to be taken into consideration. Therefore, raters needed to start with point '5' on each rating scale, and move down the scale as more alliance elements were missing. For example, when rating "understanding" between the client and the therapist, the rater should move one point lower on the rating scale for any indication of misunderstanding, such as the therapist incorrectly reformulating the client's speech. Similarly, the rater needed to provide a lower rating if insufficient understanding was communicated through mechanical repetition of the client

wording. A high score meant a more intense and better alliance component. Guidelines for each specific item were outlined. For example, Figure 6.1 presents the guidelines for rating Item 9 of the WAI-SR-O. For this item, 'Mutual Liking' is defined and guidelines on what to look for when rating mutual liking between the therapist and the client are suggested. Statements for each anchor point of the rating scale are outlined with examples to facilitate the rating of the frequency and intensity levels (e.g., Optimum, Satisfactory, Moderate, Poor and Extremely Poor).

9. Bond Item

①	②	③	④	⑤
Seldom	**Sometimes**	**Fairly Often**	**Very Often**	**Always**

There is a <u>MUTUAL LIKING</u> between the client and therapist.

Mutual liking involves having pleasant and warm interactions, being engaged, being sensitive, and showing personal interest in each other (e.g., the therapist should enhance or reinforce client expressions of personal values or outside interests, and the therapist should not forget important details about the client's life). There are open expressions of rapport, empathy, warmth, concern, genuineness and humour. Body language is open and relaxed and there are warm tones in the client's and therapist's voice. **In this item the emphasis is on mutual and refers to feelings and experiences between the client and therapist.**

Rating the intensity of mutual liking between the therapist and client can be determined by positive, warm versus negative, cold, hostile interactions, sensitivity, empathy, rapport, open and relaxed versus closed off body language.

5. There are frequent instances of mutual liking between the therapist and client or no explicit need for mutual liking to be present (e.g., if there is no pain expressed by the client). There is an **OPTIMAL** level of warmth, sensitivity, empathy, rapport and open and relaxed body language. There are warm tones in the clients and therapists voices. Any ruptures/strains are immediately investigated and repaired.

4. There are very often instances of mutual liking between the therapist and client. However, there was **ONE** instance when mutual liking may have been appropriate, but was not expressed. There is a **SATISFACTORY** level of warmth, sensitivity, empathy, rapport and open and relaxed body language. Any ruptures/strains are investigated and every effort is made to repair them.

3. There are fairly often instances of mutual liking between the therapist and client. However, there is **TWO** instances when mutual liking was clearly needed, but unexpressed. (e.g., the therapist focuses on the task at hand without being attuned to the clients need for validation). There is a **MODERATE** level of warmth, sensitivity, empathy, rapport and open and relaxed body language. The relationship may seem strained, at times but generally every effort is made to investigate the problem and repair it.

2. There are sometimes instances of mutual liking between the therapist and client. However, there are **THREE** instances where mutual liking was clearly absent (e.g., therapist's voice is mechanical and abrupt). Interactions between therapist and client are **POOR,** but do not seem destructive. Some effort is made to repair ruptures/strain.

1. There are seldom instances of mutual liking between the therapist and client. There are **FOUR OR MORE** instances when the unaddressed need was more blatant and mutual liking didn't happen as often as it should. Interactions are **EXTREMELY POOR** and destructive. Interactions seem negative, cold, and hostile. No effort is made to repair ruptures/strains.

Figure 6.1: Guidelines for rating Item 9 of the WAI-SR-O.

<u>Rating the WAI-SR-O in the Current Research</u>

Two female postgraduate psychology students were employed as independent raters and trained in the use of rating the WAI-SR-O with the guidelines (see Chapter 7 for training details). They received payment for their duties as a rater. Neither rater had previous rater training. One rater had completed some CBT courses/workshops and had work experience in the mental health field, whereas the other rater had no CBT or work experience in this field. Originally four raters were trained to rate the WAI-SR-O, but two pulled out due to their academic commitments. Training procedures ensured that they gained a thorough knowledge of the alliance construct and of scale items and scoring procedures of the WAI-SR-O (Hill & Lambert, 2004). Based on prior research (Gaston et al., 1998), this training continued until the reliability between the raters was acceptable. They rated the alliance of therapy sessions 1-10 of each participant in random order to prevent drift (Gaston et al., 1998; Kazdin, 1977; Kobak, Williams, & Lipsitz, 2008). Ratings were then made independently after watching an entire session of therapy and then the scores were averaged.

Recalibration took place after the completion of 10 consecutive ratings. See the results section for further description of the ICCs for overall recalibration of 10, 20, and 30 consecutive ratings and for ICCs of the recalibration of items for 10, 20, and 30 consecutive ratings. It was planned that recalibrations would be conducted after each 10 consecutive sessions to assist retraining requirements. However, this plan was not adhered to after the third set of ten consecutive ratings due to the current researcher's internship commitments. After each recalibration feedback was given to the raters and any items with low reliability were discussed. For example, reliability was lower for the second set of ten ratings. Raters reported that in some of these sessions, the therapist-client interaction was more difficult to hear due to a crying baby and a thunderstorm recorded onto the DVD or the client was speaking too quietly. These factors may have affected the reliability for this set of ratings. Furthermore, they had noticed that they became fatigued by the third session if they tried rating three sessions on the same day. What the raters focused on or put emphasis on within each specific item, levels of agreement/collaboration and engagement and scoring difficulties were discussed during feedback. It should also be noted that not only were the raters rating the WAI-SR-O, they were also concurrently rating observer versions of the HRS-II, the HAACS and the Conceptualization Rating Scale. However, this concurrence of rating scales alongside each other is common. (R. J. DeRubeis, personal communication, 26th October, 2007). The overall inter rater reliability (ICC) for the

WAI-SR-O for the current research was 0.68 with 95 % CI from 0.65 to 0.71. The ICC for the Goal subscale was 0.74, for the Task subscale the ICC was 0.54 and for the Bond subscale the ICC was 0.64. These results are reviewed in the results section (see Chapter 8).

Ethical Considerations

The CBT for Depression Outcome Study, conducted at *The Centre for Psychology*, School of Psychology, Massey University, Albany was designed in accordance with the ethical guidelines of the New Zealand Psychological Society and has been approved by the Northern X Regional Ethics Committee (NTX/06/07/085 - The Relationship Between Therapist Competence in Using Homework Assignments, Patient Homework Adherence, and Treatment Outcome in Cognitive Behaviour Therapy for Depression). Ethical issues considered in this study included: privacy and confidentiality of participant information, informed consent of the participants once they had gained and thoroughly understood all the information about the study, therapist training and supervision, the safety monitoring of individual participants and the overall study to ensure that the participants were not exposed to undue risk, and cultural responsibility. Participants were informed at the onset that the study did not have provisions to offer a culturally specific/adapted service. They could then make an informed decision as to whether they wanted to participate in standard CBT and if individuals preferred not to be involved in the study, they could be referred onto other services where treatment was specifically designed for their culture (e.g., Maori Mental Health Services at Waitemata District Health Board, MOKO services). For Maori who elected to be involved with the study, a kaumatua, koro Turoa was available for guidance on cultural issues. Two Maori lecturers at the School of Psychology were consulted about the projects cultural responsibility.

Research Design of the Current Research

The current research employed a single-case research design to investigate the role of the alliance and symptomatic change within CBT for depression. It utilized the archived session data recorded on digital video discs (DVDs) from the Depression Outcome Study to rate the alliance process of the first ten sessions of therapy of the first ten client-therapist dyads in the study. These ten clients completed the Beck Depression Inventory–II (BDI-II; A. T., Beck et al., 1996) at the beginning of every therapy session and this data was used in the current research to assess symptomatic change. Data was displayed graphically (i.e., simple line graphs) to visually examine patterns of alliance scores; the overall WAI-SR-O

score and its subscales (i.e., Goal, Task and Bond) and symptomatic change (i.e., BDI-II) for each case.

Single-Case Research Design

The single-case research design has generated significant debate in the social sciences (Mahoney, 2000). It has been typically discounted as a potential source of scientifically validated inferences because threats to internal validity can not be ruled out in the manner of that achieved by experimentation (Barlow & Hersen, 1984; Hersen & Barlow, 1976; Kazdin, 1998, 2001, 2003, 2010). All research designs must permit reliable changes in the dependent variable to be detected despite variance and they must permit valid inference to be drawn as to the cause of any changes observed (Barlow & Hersen, 1984; Blampied, 1999, 2001; Hersen & Barlow, 1976; Kazdin, 1998, 2001, 2003). Single-case research designs achieve these objectives by maintaining commitment to quantification, involve rigorous (but different) experimental designs and use replication as its key justification procedure (Barlow & Hersen, 1984; Sidman, 1960). The single-case research design demonstrates efficacy and the effectiveness of therapy (Chambless & Hollon, 1998) and they are ethical in that once treatment begins, it is not necessary to withdraw or alter the intervention to demonstrate the treatment effects (Kazdin, 1998). Therefore, they are user friendly in clinical settings. Single-case research designs avoid problems inherent in the averaging of data over many individuals, or when data is collapsed by averaging over time (Blampied, 2001). They are flexible and are not limited to proving or disproving a null hypothesis. Single-case research designs permit science to be conducted at the level of the individual and variability of human behaviour can be captured in response to therapeutic intervention (Barlow & Hersen, 1984; Blampied, 2001; Morgan & Morgan, 2003). Single-case research designs provide a rich source of ideas for developing hypotheses for future research (Kazadin, 2003, 2010).

 Replication

One of the most frequently mentioned limitations of single-case research designs is their presumably minimal external validity in terms of generality (Morgan & Morgan, 2003). However, Kazdin (2003) argued that the external validity of single-case research depended on systematic replications of effects in many clients. It may take two forms in different single case rsearch designs: within person replication and between person replication. For any claim to valid causal inference, one complete replication is required, but the more replications available, the stronger the validity of the causal inference can be made (Barlow

& Hersen, 1984; Blampied, 1999, 2001; Cooper, Heron, & Heward, 1987; Hersen & Barlow, 1976; Kazdin, 1998, 2003, 2010; Schneider, 2001). Conclusions made from several individuals can rule out the possibility of idiosyncratic findings that ae characteristic of one individual (Blampied, 1999, 2001; Hayes, 1998; Kazdin, 1998, 2003, 2010).

Repeated Measurement

Standard single-case research designs involve making repeated assessment of an individual or group (Blampied, 2001, 2003; Hayes, 1998; Kazdin 1998). This continuous assessment strengthens the internal validity of the study (Kazdin, 1998; Morgan & Morgan, 2003). The changes that coincide with treatment are not likely to result from exposure to repeated testing or changes in the instrument (Kazdin, 1998, 2003). Regression to the mean from one data point to another, a special problem with assessment conducted only at two points in time is eliminated. Repeated measurement over time shows a pattern in the data.

Visual Analysis

There has been little consensus on statistical methods for analyzing sequence response data from single cases (Morgan & Morgan, 2003, Kazdin, 2010). Visual inspection still dominates single case research designs. Graphical display (e.g., simple line graph) is particularly useful for seeing the patterns in the data obtained over time (Kazdin, 2003, 2010). Single-case researchers justify graphic display by maintaining that the evidence of an independent variable's effect on the dependent variable should be visible to the naked eye (Kazdin, 2003, 2010; Parsonson & Baer, 1978). Continuous data representing the level of response for each participant can be plotted at each assessment point and changes in level, slope and trend are easily examined and interpreted at a glance (Kazdin, 2003, 2010).

Data Preparation

Raw data obtained from rating the WAI-SR-O was entered into a spreadsheet for analysis using Statistical Package for Social Sciences, Version 17.0. The raters' score for each item of the WAI- SR-O was added together and then averaged. The overall Total Alliance score was comprised of the sum of the scores for each of the 12 items. Each subscale score was comprised of the sum of the scores for each of the four items belonging to that subscale.

Missing Data

Prior to analysis the data was checked for missing values. There were five scores missing from the WAI-SR-O data due to rater error. These scores were replaced with the medium score of the rater for that particular subscale for that session.

Statistical Analysis

Inter-rater Reliability

Reliability may be defined as the consistency with which an instrument discriminates among a group of raters (Kozolowski & Hattrup, 1992). Inter-rater reliability is usually defined as the proportion of systematic variance in a set of ratings in relation to total variance of ratings (Shrout & Fleiss, 1979). While it is common to study agreement among ratings of multiple judges there is also little consensus about what statistical methods are best to analyze rater agreement (Uebersax, 2003; Vincent, 2002). Considering theory and knowing the goals when analyzing agreement data is important (Uebersax, 2003). In the current research inter-rater reliability ensured consistency of ratings made by the two raters during recalibrations, provided feedback to the raters about areas of concern and provided overall reliability of the WAI-SR-O for the current research.

There are many methods for estimating inter-rater reliability. Intraclass correlation is derived from a repeated measures analysis of variance and is the method most often used with interval data (Tinsley & Weiss, 1975). Consistent with recent literature that evaluates inter-rater reliability within CBT (Barber, Liese, & Abrams, 2003; Persons & Bertagonolia, 1999) Intraclass Correlation Coefficients (ICC; Haggard, 1958; Shrout & Fleiss, 1979) were utilized to estimate inter-rater reliability in the current research. Data was analysed using the Statistical Package for Social Sciences, Version 17.0.

Based on the use of two or more independent judges, ICCs provide reliability estimates that account for chance and systematic differences between raters (Shrout & Fleiss, 1979) and are therefore more appropriate than Pearson's product movement, Spearman's correlations or percent agreement. Furthermore, ICCs are preferred over Pearson's r when the sample size is small (< 15) and Pearson's r does not make any assumptions about rater means (Walter, Eliasziw, & Donner, 1998). ICCs assess rating reliability by comparing the variability of the different ratings of the same subject to the total variation across all ratings and all subjects (Fleiss, 1973, 1981; Shrout & Fleiss, 1979). Large positive values approaching 1.0 occur when there is no variance due to the raters and no residual variance

to explain. When there is not sufficient variation between the cases the ICC may be low and possibly negative or uninterpretable, even when there is substantial agreement among raters (Bartko & Carpenter, 1976). ICC calculations are sensitive to variance in responses and can produce negative results when between-group variation is lower than within-group variation indicating that a third variable has introduced non-random effects on the different groups (McGraw & Wong, 1996; Rosner, 2006). Therefore the presence of low and negative ICCs do not necessarily suggest that these items are not reliable, but may indicate the presence of variability within–rater responses for low base-rate behaviours (Thompson, 2003). ICC is 0 when within-groups variance equals between-groups variance, indicative of the grouping variable having no effect (Rosner, 2006). Established guidelines evaluating ICC levels suggest that a coefficients of .80 and above be regarded as excellent, .60 to .80 as satisfactory, .40 to .60 as acceptable but possibly improvable, .20 to .40 as demanding improvement and .20 or below as totally unacceptable (Landis & Koch, 1977; Portney &Watkins, 1993; Rosner, 2006).

There are three types of ICC with the choice depending on the purpose of the study, the research design and the measurement used (McGraw & Wong, 1996; Shrout & Fleiss, 1979). In the current research the raters represented a sample from the population of all possible raters. Therefore, Model 2 was utilised for the current research. Model 2 is generally chosen to show that a measure has broad applications (Portney & Watkins, 1993; Shrout & Fleiss, 1979). Three other subtypes of each ICC type are used. First, the analysis may be a one-way or two-way design depending on whether one or both of the scores (rows) and the raters (columns) are being assessed (Shrout & Fleiss, 1979). The current research utilized a two-way design because the raters and the scores were both deemed important factors in the ICC computation. Second, there are two forms for the ICC, depending on whether the study assesses the single scores of any one rater or uses the average scores of a group of raters (Shrout & Fleiss, 1979). The current research was interested in the averages of the ratings for each item, so the average measure reliability option (i.e., average reliability of the mean of the ratings of all the raters) was chosen. Finally, there are subtypes of the ICC depending on whether the design considers a systematic variation between the raters to be of importance. The consistency version of the ICC ignores such a relationship; the absolute agreement version accounts for it (McGraw & Wong, 1996; Portney & Watkins, 1993; Shrout & Fleiss, 1979). The assessment of the variation between the raters was an objective of the current research, and the absolute agreement was therefore required. The Shrout and Fleiss (1979) ICC (2,k) indicates a two-

way, average measure calculation, where *k* is the number of raters. Therefore the model utilized in the current research was ICC (2,2).

With interval-level data various causes may result in rater disagreement on a given case. These effects can be broadly grouped into three categories: effects on the correlation or association of raters' ratings, rater bias, and rater differences in the distribution of ratings (Uebersax, 2003).

Rater Association

Association concerns whether raters understand the meaning of the trait in the same way (Fleiss, 1973). In making a rating, raters typically consider many factors. For example, in rating alliance, a rater may consider separate factors of frequency of alliance behavior that was helpful or clearly unhelpful and the intensity of the interactions. Judgments on these separate factors are combined by the rater to produce a single overall rating (Uebersax, 2003). Raters may vary in what factors they consider, weight the same factors differently, or they may use different algorithms to combine information on each factor to produce a final rating. Finally random effects such as raters being subject to distractions or the focus of the rater varying may affect the rating process (Uebersax, 2003). The extent to which raters' ratings correlate less than 1, is evidence that the raters are considering or weighting different factors in the rating process (Uebersax, 2003). When rater association is low, it implies that training methods to improve the consistency of raters' criteria need to be applied (Bartko & Carpenter, 1976). Feedback may then promote more consistent ratings. ICCs take into account the association as well as the level difference between observer ratings (Bartko & Carpenter, 1976).

Rater Bias

Rater bias refers to the tendency of a rater to make ratings that are higher or lower than those of other raters. Bias may occur for several reasons (Kim, 1995; Rudner, 1992). For example, some raters may tend to over rate alliance or under rate the alliance or they may simply interpret the calibration of the rating scale differently. Rater bias can be assessed by calculating the mean rating of a rater across all cases that they rate. High or low means, relative to the mean of all raters, indicate positive or negative rater bias, respectively (Uebersax, 2003). ICCs are sensitive to rater mean bias (Uebersax, 2003).

Rater Distribution

Sometimes an individual rater will have a noticeably different distribution than the distribution of ratings for all raters combined (Uebersax, 2003). In the current research this may partly relate to rater differences in what they believe is the strength of the alliance within the therapist-client dyad. Examination of the distribution of ratings by each rater may sometimes reveal important differences. Graphically displaying the distribution of each rater's ratings, and the overall distribution, and the base comparisons on these displays is a useful way of calculating rater distribution (Uebersax, 2003).

Chapter 7 outlines the training for the raters in the current and research and investigates whether giving more information (i.e., guidelines developed for rating the WAI-SR-O) leads to higher inter-rater reliability.

Overview

It has been reported that rating the alliance can be extremely subjective (L. R. Fenton, personal communication, 1st September, 2007; R. L. Hatcher, personal communication, 19th April, 2007; Horvath, 1994a; Horvath & Greenberg, 1994; Raue et al., 1993; Raue et al., 1991). As a construct, the alliance is complex and difficult to define (Andrusyna, et al., 2001; Bordon; Gaston, 1990; Gelso &Carter, 1985, 1994; Luborsky, 1994) and agreement on scale items can involve much debate (L. R. Fenton, personal communication, 1st September, 2007; Raue et al., 1991). There is doubt whether any brief general alliance scale can detect collaborative processes between therapist and client within CBT (Hatcher, 2006). Therefore, to facilitate high inter-rater reliability, the current research modified and extended existing guidelines for rating the observer version of the original WAI (Raue et al., 1991) to provide more specific criteria for the WAI-SR-O, in order to capture and rate the alliance within CBT. This chapter details rater training of the current research and an experimental study conducted within the current research to establish inter-rater reliability. Figure 7.1 demonstrates how Chapter 7 fits into the current research. This rater reliability study was carried out in two stages. Stage One investigated whether using the guidelines to rate the WAI-SR-O leads to higher inter-rater reliability. Stage Two was conducted to test the findings of Stage One. It investigated whether there was higher inter-rater reliability rating the alliance within CBT if raters had work experience in the mental health field or a related practice and/or CBT training experience. Chapter 7 begins with a brief outline of the methodological issues surrounding rating errors, rater training and retraining and the needed characteristics of raters to reduce error and obtain adequate inter-rater reliability. The methods and the results of Stage One and Two are presented, followed by a discussion of the results of both stages.
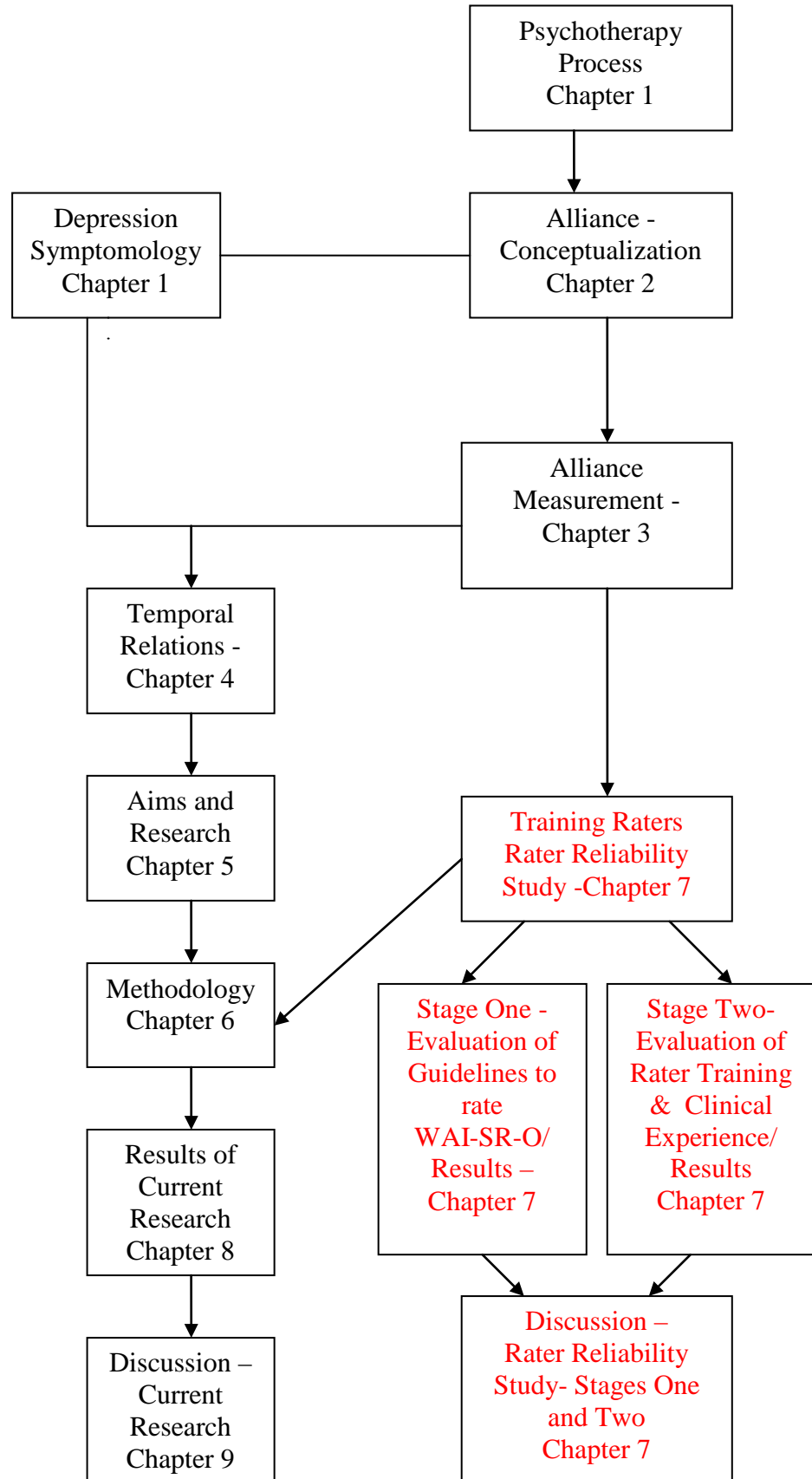
Figure 7.1: Thesis outline demonstrating how Chapter 7 fits into the current research.

Methodological Issues Surrounding Rater Training.

Rating in psychotherapy is often difficult, but estimates of consistency of agreement between raters have become increasingly important. High quality ratings are essential for reliable and valid ratings of therapy processes (Rudner, 1992). Rating errors can undermine the reliability of the information provided by the raters (Bannister, Kinicki, Denisi, & Horn, 1987). Quality of ratings can also be determined by rater training (Athey & McIntyre, 1987; Bernadin & Pence, 1980; Kobak et al., 2008; Shohamy et al., 1992; Reichelt, James, & Blackburn, 2003; Latham et al., 1975) and retraining (Ivancevich, 1979; Warmke & Billings, 1979), and rater characteristics (Hambleton & Powell, 1983; Heneman, Schwab, Huett, & Ford, 1974; Holzbach, 1978; Zedeck & Cascio, 1982).

Rating Errors and Biases

Rating errors occur when one individual observes another and is unable to make totally observable judgments (Thorndike, 1920).Traditionally the quality of performance ratings has been evaluated in terms of rating errors such as halo error, leniency/severity error, central tendency error, restriction of range error and inter-rater discrepancy error (Hill & Lambert, 2004; Saal, Downey, & Lahey, 1980). Halo error refers to the extent to which a rater is likely to give similar ratings to all performance scores by generalizing overall impression about the ratee/stimulus (Nisbett & Wilson, 1977; Tsui, 1983). Leniency/severity error refers to the tendency on the part of the raters to consistently provide ratings that are higher/lower than is warranted by the performance (McIntype, Smith, & Hassett, 1984; Saal & Landy, 1977). Central tendency error refers to the tendency to rate all rating objects around the middle or mean of the rating scale and avoid the use of the extremes (Korman, 1977). Restriction of range error refers to the extent to which one rater has a tendency to give similar ratings to all ratees and/or items using a narrow portion of the rating scale; reluctance to use the entire scale range when assigning ratings (Vance, Winne, & Wright, 1983). Inter-rater discrepancy error refers to the extent to which two or more raters, working independently, disagree on which phenomena occur to what degree in the target of interest (Freeberg, 1969; Tsui, 1983). However, when examining ways to improve the quality of ratings Kim (1995) found no consensus for definitions of each rating error. This author also found that some of the errors were used differently in different research and frequently only one source of rating error was examined within each study. Saal et al. (1980) demonstrated that different operational definitions can result in different conclusions. The scope of this thesis does not permit further discussion on this topic, but there is a multitude of literature surrounding rating

criteria. However, virtually all this research is found in the field of applied psychology with studies being conducted in educational (Bernardin & Walter, 1977; Freeberg, 1969; Holzbach, 1978; Kim, 1995; McIntyre et al., 1984) and organizational (Borman, 1975, 1979; Ivancevich, 1979; Saal & Landy, 1977; Warmke & Billings, 1979) settings. Kim's (1995) review provides insight into the problems and advantages associated with the different definitions in rating errors.

Training and Retraining Raters

Raters must be trained to successfully carry out their roles in performance ratings (Borman, 1979; Reichelt et al., 2003; Latham et al., 1975). The goal of training raters in the current research was is to build rater knowledge and assist them to become familiar with the scoring criteria of the WAI-SR-O. Over the last three decades much effort has been devoted to the analysis of alternative rating techniques for improving rater quality (Bernardin, 1977; DeCotiis, 1977; Keaveany & McGann, 1975; King, Hunter, & Schmidt, 1980; Saal & Landy, 1977) and to evaluating the usefulness of training for reducing rater errors (Bernardin, 1978; Borman, 1979; Latham et al., 1975). Several studies have shown that intensive error training can reduce errors, particularly if rating errors are discussed in the early stages of training (Bernardin, 1978; Bernardin & Walter, 1977; Borman, 1975; Ivancevich, 1979; Latham et al., 1975). To yield more reliable and accurate ratings it has been demonstrated that training needs to focus on how to interpret the data, and how to use the data to formulate judgments (Pulakos, 1986). Raters need clear understanding of procedures used to rate or evaluate scale items and the criteria for determining the quality of performance (Bernardin et al., 1976; Kiesler, 1973; Kobak et al., 2008; Nott, Reeve, & Reeve, 1992). The use of structured training sessions (Reichelt et al., 2003), rating stimulations, such as video tapes and role plays (Borman, 1979; Jaeger & Busch, 1984), and group discussion during training (Levine & Butler, 1952) have been found to contribute to successful rater training. Furthermore, practice ratings in a training programme enable raters to receive feedback, thereby assisting them to recognize their rating patterns and improve in certain areas where reliability is lower (Hill & Stephany, 1990; Kim, 1995; Kobak et al., 2008).

The few studies that have examined the effects of training on rating over time showed that rater training effects dissipate after 6-12 months indicating the need to provide reinforcement and refresher training (Bernardin, 1978; Ivancevich, 1979; Warmke & Billings, 1979). There is also a tendency for raters to change the manner in which they

apply the definitions of the behavior (Kazdin, 1977). As familiarity increases and routines become established raters may make more assumptions and snap decisions about the data (Hill & Lambert, 2004). Raters may also be more accurate when they think their behaviour is being checked or monitored (Mitchell, 1979; Reid, 1970; Romanczyk, Kent, Diament, & O'Leary, 1973). Kim (1995) found that the best results were obtained when feedback was continuous, consistent, and conducted on an informal basis.

In conclusion, there a is a multitude of studies demonstrating that rating errors may be reduced by giving raters an orientation about the concepts of the project, defining the major purposes for rating, giving perspectives on ratings, clarifying various measurement errors, providing guidelines to avoid common rating errors, demonstrating the use of scoring rubrics properly, providing practice ratings and feedback and allowing raters the opportunity to raise questions regarding interpretation of the scoring system. Continuous feedback provides reinforcement and the maintenance of training effects. However, the generalisability of research surrounding the training and retraining of raters may have been limited due to some studies not having a control group and therefore observed effects may be due to practice with the scales rather than the training (Borman, 1975), the artificiality of the rating task (e.g., rating vignettes about individual hypothesized first-line supervisor ratees) (Borman 1975), small samples (Warmke & Billings, 1979), the sample of participants being employed in specific jobs (e.g., supervisory engineers) (Ivancevich, 1979) and a special sample of raters (e.g. 156 students of 13 different instructors of a general psychology course) (Bernardin & Walter, 1977). Furthermore, much of the research studying the training of raters in order to reduce errors was conducted pre mid-1980s, before process researchers changed their focus from therapy process or describing specific occurrences within therapy to the process of change or the effect of in-session behaviour on eventual client outcome (Pachankis & Goldfried, 2007). Therefore, while these studies provided a solid foundation for training raters in the current research, it was difficult to determine how much of these results could be generalized to training raters to rate psychotherapeutic processes, such as the alliance.

Characteristics of Raters

Ratings may also be influenced by the characteristics of raters (Hambleton & Powell, 1983; Kiesler, 1973; Kim, 1995; London & Poplawski, 1976; Moritsch & Suter, 1988; Moras & Hill, 1991; von Raffler-Engel, 1980). Klien, Mathieu-Coughlan and Kiesler (1986) proposed that time spent training raters and the actual work of raters can be costly and

therefore it is important to identify variables that may be related to good rating performance. Hambleton and Powell (1983) argued that demographic variables such as race, age, gender, education, occupations, specialty and willingness to participate should be considered in the selection of raters. However, the results of studies examining these factors are mixed. For example, von Raffler-Engel (1980) found that professional women were consistently harder raters than men, London and Poplawski (1976) demonstrated that women were more lenient in their ratings, but Moritsch and Suter (1988) found that raters rating error patterns were similar regardless of rater gender. However, the generalisability of this research was limited due to small sample size (e.g., von Raffler-Engel, 1980), student only samples (Moritsch &Suter, 1988), laboratory interviews rather than real interviews (Raza & Carpenter, 1987) and that the studies were only conducted in organizational (London & Poplawski, 1976; Raza & Carpenter, 1987; von Raffler-Engel, 1980) or educational settings (Moritsch & Suter, 1988).

More recently Kim (1995) demonstrated that while age may not directly influence rating errors, as raters' age increased less rating errors were committed. After examining the relationship between the number of ratings (i.e., a task of reading a proposal application for school funding and then rating the proposal on 25 items) and the influence on rating errors this author found that as the number of ratings increased for each rater, rating errors increased. Raters who rated more than seven ratings often committed errors. This result was interpreted as fatigue and/or boredom affecting rating errors. However, the author did not stipulate whether these ratings were conducted in one sitting or over a period of time. Younger raters rating patterns were more inconsistent after seven ratings and those older than 50 years committed more rating errors after seven ratings. However, as the number of ratings exceeded seven raters aged 40 - 49 years committed relatively fewer errors.

Raters may also differ with their life experience, specialty and continuing professional skills. However, research findings are inconsistent in this area. Research studying the effect of the rater's (i.e., English teachers versus lay persons) background and training on the reliability of direct writing tests suggested that raters are capable of rating reliability, regardless of background and training (Shohamy et al., 1992). However, Kim (1995) demonstrated that selecting raters who are knowledgeable about the topic being rated (i.e., teachers with 15 and more years teaching experience utilizing a 25 item education instrument to rate proposal applications for funding) committed the least rating errors. Persons and Bertagonolia (1999) found that therapists with PhD level training were more

accurate than non PhD therapists in identifying patient problems. Furthermore, Kim (1995) demonstrated that raters with previous rater training and rater experience committed more leniency, central tendency and inter-rater discrepancy errors, but less halo, severity and restriction of range errors than those who had no previous rating experience.

Research surrounding selection criteria for raters in process research is sparse and most literature in this area is founded on research conducted in organizational and educational psychology. After reviewing prominent psychotherapy observer-rated process measures and available published rater selection information Moras and Hill (1991) found that the level of clinical experience needed by raters was directly related to the amount of inference required by the task. Arnhoff (1954) concluded that trained, clinically naïve undergraduate raters could rate highly operationalized variables such as nonverbal behavior because they are more likely to attend to what they have been training to observe and not be biased by clinical experience, whereas inter-rater agreement decreased with increasing clinical experience. Shapiro (1968) found that the ratings of untrained undergraduate psychology students using undefined scales of psychotherapy behaviour (e.g., empathy, warmth and genuineness) showed a significant, positive correlation with the ratings of professional trained people. Clinically experienced clinicians were used in a variety of counseling research studies (Burstein & Carkhuff, 1968; Carkhuff, Kratochvil, & Friel, 1968; McMullin, 1972; Mitchell & Brenson, 1970) and it was suggested that only high-level functioning and clinically sophisticated persons could validly discriminate highly diverse levels of functioning (Burstein & Carkhuff, 1968). However, empirical evidence is inconclusive and may still depend on how the alliance construct is defined.

Motivation can be a major factor in determining rating error (Feldman, 1978; McIntype et al., 1984; Moritsch & Suter, 1988) and it has been suggested that raters need to be motivated to participate in ratings to provide quality ratings (Feldman, 1978; Kim, 1995). Kim (1995) found that raters whose motivation was high committed less halo, severity, central tendency and restriction of range errors, but committed more errors in leniency and inter-rater discrepancy errors Therefore, the task needs to be meaningful (Hill & Lambert, 2004).

In conclusion the choice of raters may have a significant impact on the quality of ratings, but there is a lack of research in the area of selecting raters for process-outcome research. It has been suggested that considering demographic variables, choosing a motivated and

knowledgeable rater with no previous rater training and rater experience may improve ratings. However, there is not enough evidence to suggest that selecting raters according to previous rater training and experience is a reliable selection criteria. It has also been proposed that experienced clinicians are needed to rate complex processes such as the alliance, but more research is needed in this area. Furthermore, process-outcome researchers utilizing observer versions of alliance scales rarely documented rater selection, characteristics and training; therefore successful replication of these clinical studies may not have been possible. This may also have contributed to the inconsistency of results in this area of research.

Rater Reliability Study

Research Objectives - Stage One

Aim and Hypothesis

It has been suggested that giving clear guidelines regarding procedures to rate or evaluate scale items can increase rater reliability (Bernardin et al., 1976; Nott et al., 1992). The aim of Stage One was to investigate whether giving more information (i.e., training participants with guidelines developed for rating the WAI-SR-O) would lead to less variance in ratings and higher inter-rater reliability. It was therefore hypothesized that training raters with additional guidelines developed for rating the WAI-SR-O would lead to greater inter-rater reliability.

This current study was interested in investigating not only the with-in and between group differences, but also the differences between the two groups and expert raters when utilizing the guidelines. It was expected that the inter rater reliability of the experts would have been high and would have set the 'gold standard' of rating the WAI-SR-O with the guidelines. Unfortunately this could not be explored further due to more expert raters not being available at the time of the present study.

Method

Participants

Eight participants (i.e., 6 female and 2 male with ages ranging from 22 -52 years old) were recruited as raters through an email sent out to psychology postgraduates. As each participant replied to the email they were randomly assigned to two groups of four. The focus of study for most of the participants was clinical (one had studied health psychology) and four had working experience in the mental health field or a related practice. This work experience (e.g., community mental health support worker, caregiver in dementia ward, disability mentor for university students, telephone counselling and phone support lines for mental health) ranged from six weeks to eight years. All the participants were Caucasian. Table 7.1 presents participant demographic characteristics of the participants in Group A and Group B. Each participant signed a Massey University Centre for Psychology confidentiality form (Appendix D). All participants received payment for participating in the study.

Table 7.1: Demographic Characteristics of Participants in Group A and Group B

Group A

Participant	Age	Gender	Focus of post graduate study	CBT Training Experience	Work Experience
1	50 yrs	Female	Health	None	None
2	30 yrs	Female	Clinical	Psychotherapy paper at Massey University Various CBT workshops at NZPS conference	8 years
3	22 yrs	Female	Clinical	None	None
4	47yrs	Male	Clinical	None	None

Group B

Participant	Age	Gender	Focus of post graduate study	CBT Training Experience	Work Experience
5	52yrs	Female	Clinical	Psychotherapy I[1]	2 years 6 months
6	24yrs	Female	Clinical	Psychotherapy paper at Massey University Padesky and Mooney's Workshop -resilience	2 years 4 months
7	48yrs	Female	Clinical	Psychotherapy paper at Massey University	None
8	23yrs	Male	Clinical	Psychotherapy paper at Massey University	6 weeks

Measure

WAI-SR-O and Guidelines

The WAI-SR-O and the newly developed guidelines for rating the WAI-SR-O were utilized for this small experiment. The WAI-SR-O and the guidelines for rating the WAI-SR-O were described in the previous chapter.

---

[1] The postgraduate paper "Psychotherapy I: Theory, Research and Practice" contains an introduction to CBT principles and structure, including CBT for depression.

Procedure

Rater Training

Initial Training

Initial training for rating the alliance took place over two consecutive eight hour days (see Appendix E for powerpoint presentation of training). It was conducted by the current researcher (CO) and her supervisor (NK). On the first day participants in both Group A and Group B received a brief overview of the current research within The Depression Outcome Study. It was important that raters felt that what they were doing was important and that their contributions toward the project were valued. This overview also emphasized the importance of defining the term alliance within CBT and finding the most appropriate instrument to operationalize the alliance concept within CBT. All participants were then given the opportunity to brainstorm their existing knowledge of the alliance. They broke into pairs to write down there ideas and then this information was displayed and presented back to the whole group. The idea that most theoretical definitions of the alliance have three common themes (i.e. the collaborative nature of the relationship, the affective bond between the client and therapists and the client and therapists ability to agree and work on treatment goals and tasks) was then elaborated. It was explained that within CBT therapist characteristics such as warmth, empathy, and respect were necessary to engage clients in the process of collaboration empiricism, but that these factors were not sufficient to produce optimum therapeutic effect. Collaboration empiricism was defined and emphasized as encapsulating the idea that the client and therapist work as a team. Existing alliance measures and the development of the WAI-SR-O were briefly discussed. Finally concepts based on Bordin's model of the alliance and Goal, Task, and Bond dimensions of the alliance were thoroughly explained. Participants were shown how items of the WAI-SR-O were broken down into Goal, Task and Bond subscales and each item was discussed. To maximize understanding, instances of the Goal, Task and Bond subscales were demonstrated from recorded therapy sessions from the Depression Outcome Study and then related back to items of the WAI-SR-O. Discussion followed to ensure that all raters understood the concepts.

Guideline Training

In the guideline training each point of the rating guidelines and the guidelines for scoring each specific item were discussed and any questions answered. Rating the WAI-SR-O by starting with the point "5" on the scale and then moving down the scale as more and more of the alliance elements are missed and/or deemed unhelpful to the relationship was talked

about until participants clearly understood this method of evaluating frequency and intensity of the affective bond and the ability of the therapist and the client to work together.

Sessions Utilized During Rater Reliability Study

Recorded therapy sessions archived from Depression Outcome Study were utilized during this experimental study. Session difference can be a confound as therapist and client behaviour can change from the first third to the final part of the treatment (Hill, Carter, & O'Farrell, 1983; Hill & O'Grady, 1985; O'Farrell, Hill, & Patton, 1986). Therefore, therapy sessions 15-19 from the same client/therapist dyad were chosen. Each session was approximately 50 to 60 minutes long.

Experimental Design

During the first stage of the experiment participants participated in a series of three ratings of the recorded therapy sessions archived from Depression Outcome Study (see Table 7.2 for a summary of the Experiment Design). On the first day of training Group A and Group B rated Session One with the WAI-SR-O scale without guidelines. This step in the procedure allowed the difference in reliability between the two groups to be examined. The following day all participants received an expert's ratings and reasons for rating each item to reflect on and compare to their own ratings. The expert rater (5 years experience as a clinical psychologist and 3 years as a CBT therapist) had previously rated Session One with the newly modified WAI-SR-O guidelines. These ratings and reasons for rating each item provided apprenticeship for raters in training and ensured that participants were more familiar with rating alliance within CBT (Adler, et al., 2005).This step also ensured that both groups were on the same baseline before commencing the next stage of training. Further discussion was encouraged to clarify concepts and any misunderstandings.

On the second day of training participants in Group A were given additional training with the newly developed guidelines for rating the WAI-SR-O. Group A and Group B rated Session Two. Group A used the rating guidelines and WAI-SR-O scale, whereas Group B continued to use the WAI-SR-O scale without guidelines. This step demonstrated the difference in reliability between the two groups for Session Two and the change in reliability between Session One and Two for Group A after the additional training. It also enabled improvement above the practice effect to be examined in Group B by showing the changes in reliability between Session One and Two.

Participants in Group B were then trained with the guidelines for rating the WAI-SR-O. Both groups then rated Session Three using the WAI-SR-O scale and guidelines. This step in the procedure showed the change in reliability between Session Two and Three for Group B. Practice effects were examined by looking at the changes in reliability between Session Two and Three for Group A and changes in reliability between Session One, Two and Three for both Groups A and Group B. Discussion was encouraged after rating each session and each group gave feedback on any item difficulties after rating Session Three.

Table 7.2: Summary of Rater Reliability Study Design

Session	Group A	Group B
One	WAI-SR- O	WAI-SR-O
Two	WAI-SR-O and Guidelines	WAI-SR-O
Three	WAI-SR-O and Guidelines	WAI-SR-O and Guidelines

Statistical Analysis

Inter-rater Reliability
The current research was interested in the consistency of ratings made by different raters of the same therapist/client dyad to validate the use of guidelines for rating the WAI-SR-O. The ultimate goal was to improve the inter-rater reliability when rating the alliance. The effects of the between groups and within group factors during Stage One were calculated and analyzed utilizing Intraclass Correlation Coefficients (ICC; Haggard, 1958; Shrout & Fleiss, 1979) (see previous discussion on inter-rater reliability and ICCs in Chapter 9). Models used for the calculations in Phase One were ICC (2, 8) for combined group calculations (Group A & B) and ICC (2, 4) for individual group (Group A, Group B) calculations.

Results

Inter-rater Reliability of Group A and Group B during Sessions One, Two and Three
Table 7.3 presents ICCs for combined and individual groups at Sessions One, Two and Three. Overall, rater reliability improved. While the inter-rater reliability of Group A & B was 0.34 at Session One, it increased to 0.79 at Session Two and 0.81 at Session Three.

When examining group differences at Session One results indicated that Group B had higher inter-rater reliability (0.21) than Group A (0.07). In Session Two, the inter-rater reliability of Group A was 0.48 and for Group B it was 0.79. However, on closer examination of the raw data it could be seen that one of the raters in Group A had a different rating pattern than the other raters. See Figure **7.2** for a graphical display of this distribution. This participant scored only at the lower end of the range (average rating level = 1.59), whereas each of the other participants in Group A utilized most of the range of the WAI-SR-I (i.e., scores ranged between 2 and 5 with average rating level of 3.25 - 4.17). When the ICCs were recalculated without this particular rater's ratings, Group A's inter-rater reliability increased to 0.76.

Individually both groups had a decrease in inter-rater reliability from Session Two to Session Three. Group A's inter-rater reliability decreased from 0.76 (3 raters) to 0.66 and Group Bs inter-rater reliability decreased from 0.79 to 0.70.

Table 7.3: Intraclass Correlation Coefficients for Group A and B, Sessions One, Two and Three

Session	Group A & B	Group A	Group A (3 raters)	Group B
Session One	0.34	0.07		0.21
Session Two	0.79	0.48	0.76	0.79
Session Three	0.81	0.66		0.70
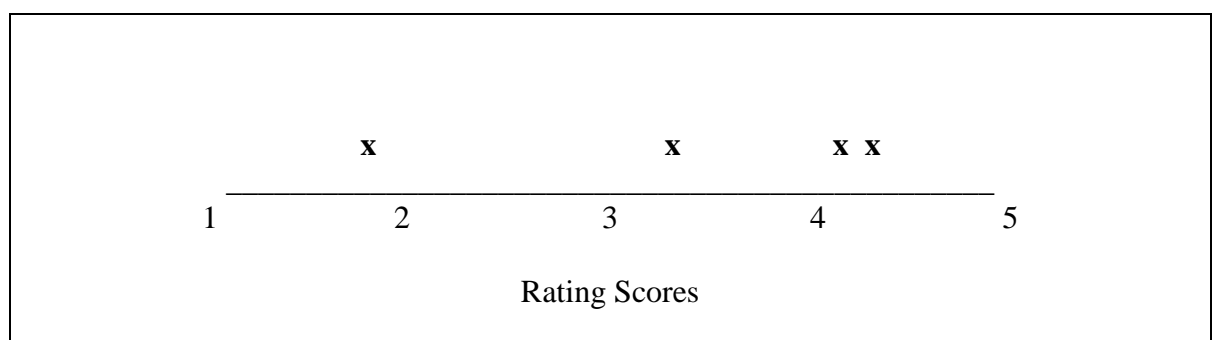
Figure 7.2: Distribution of average rating level of Group A at Session Two

Inter-rater reliability for Goal, Task and Bond Subscales

ICCs were calculated for Goal, Task and Bond subscales to explore areas of concern and to provide information regarding further training and/or improving the guidelines.Table 7.4 presents ICCs for Goal, Task and Bond subscales for combined and individual groups at Sessions One, Two and Three. Overall, inter-rater reliability for Group A & B improved for the Goal and Bond subscales. The ICC at Session One for the Goal Subscale was 0.63 and at Session Three it was 0.84. The ICC for the Bond subscale, at Session One was -0.09 and at Session Three it was 0.42. The ICC for the Task subscale was 0.25 at Session One. However, at Session Three ICC for the Task subscale was -0.58.

At Session One inter-rater reliability of Group A was higher than Group B for the Goal (Group A = 0.44 and Group B = 0.34) and Task (Group A = 0.29 and Group B = 0.07) subscales. The ICCs for the Bond subscale were negative for both groups (Group A = -0.09 and Group B = -0.24). After viewing raw data of Session One's ratings it appeared that some of the participants had rated the Bond subscale high which may indicate they were not seeing instances where the bond dimension of the alliance was strained. During feedback all participants were told that they were not as reliable on the Bond subscale before they rated Session Two.

At Session Two the ICC for the Goal subscale for Group A was 0.65 and for Group B it was 0.91. The ICC for the Task subscale was 0.00 for Group A and for Group B it was -0.21. The ICC for the Bond subscale for Group A was 0.32 and for Group B it was 0.64. After recalculating the ICCs for Group A with only three raters, inter-rater reliability for the Goal (0.89) and Bond (0.61) subscale was similar to that of Group B.

At Session Three the ICC for the Goal subscale for Group A was 0.74 and for Group B it was 0.60. The ICC for the Task subscale was -0.27 for Group A and for Group B it was -0.62. The ICC for the Bond subscale for Group A was 0.61 and for Group B it was -0.27. ICC calculations are sensitive to variance in responses and these negative results may be due to within-group variance exceeding the between group variance (McGraw & Wong, 1996).

Table 7.4: Intraclass Correlation Coefficients for Goal, Task and Bond subscales at Sessions One, Two and Three

	Goal	Task	Bond
Group A & B Session One	0.63	0.25	- 0.09
Group A Session One	0.44	0.29	- 0.09
Group B Session One	0.34	0.07	-0.24
Group A Session Two	0.65	0.00	0.32
Group A Session Two - 3 raters	0.89	0.00	0.61
Group B Session Two	0.91	-0.38	0.64
Group A & B Session Three	0.84	-0.58	0.42
Group A Session Three	0.74	- 0.27	0.61
Group B Session Three	0.60	-0.62	-0.27

Demographic Differences

After examining the above results and the demographic differences between both groups it appeared that Group B had more participants with work experience and CBT training experience. These differences warranted further exploration and participants were split into Work Experience/No Work Experience and CBT Experience and No CBT Experience groups to calculate inter-rater reliability. Table 7.5 presents a summary of which group each participant was placed.

Table 7.5: Participants in Work Experience and No Work Experience and CBT and No CBT Groups

Work Experience	No Work Experience	CBT Experience	No CBT Experience
Participant 2	Participant1	Participant 2	Participant 1
Participant 5	Participant 3	Participant 5	Participant 3
Participant 6	Participant 4	Participant 6	Participant 4
Participant 8	Participant 7	Participant 7	
		Participant 8	

Table 7.6 presents ICCs for Work Experience and No Work Experience and CBT and No CBT Groups for Sessions One, Two, and Three. When examining group differences at Session One the No Work Experience and No CBT Training groups appeared to have higher inter-rater reliability, albeit it was low for all groups. However, inter-rater reliability of the Work Experience and CBT Training groups was higher than the No Work Experience and No CBT Training groups at Session Two and Three. Inter-rater reliability decreased from Session Two to Session Three in the Work Experience (0.84 to 0.72) and CBT Training (0.87 to 0.74) groups, whereas it increased in the No Work Experience (0.42 to .60) and No CBT Training (0.23 to 0.57) groups. In Session Two (and marginally in Session Three) inter-rater reliability for those with No CBT Training was lower than those No Work Experience

Table 7.6: Intraclass Correlation Coefficients for Work Experience and No Work Experience and CBT and No CBT Groups for Sessions One, Two, and Three.

Session	Work Experience	No Work Experience	CBT Training	No CBT Training
Session One	0.07	0.22	0.07	0.20
Session Two	0.84	0.42	0.87	0.23
Session Three	0.72	0.60	0.74	0.57

Feedback from Participants about the Guidelines.

Most participants reported it easier to rate the WAI-SR-O with the guidelines. While at first participants needed to keep good notes and categorize notes, by the end of Session Three they reported taking less notes. The guidelines felt like a manual that added to the thoroughness of rating the alliance and they could look back on the guidelines when needed. They reported that counting instances was helpful. The scoring descriptions "fleshed it out", and there was more clarity to make judgments. By Session Three they were more confident and less anxious about their ratings. However, some found learning what the specific items meant; learning how to rate the items and then rate a session at the same time was difficult. The expert's ratings and reasons for rating each item was beneficial. Being "naïve" they commented that in Session One they "didn't know what they were measuring against".

The participants also feedback that some items needed teasing out more as there were too many similarities (e.g., for items 10 and 11 there needed to be more distinction between the words 'respect' and 'care', and 'mutual liking' and 'appreciation'). They also

commented that highlighting certain phrases in the descriptions of how to rate a specific item could emphasize what was required. Therefore, for item 12 "The client feels that the therapist appreciates him/her as a person", the words "sensitivity to the uniqueness of the client's plight" was highlighted and "However, this item refers to the clients perceptions" was added.

Rater Reliability Study

Research Objectives – Stage Two

Aim and Hypothesis

The aim of Stage Two was to further investigate whether participants having work experience in the mental health field or a related practice and/or CBT training experience are more reliable raters of the alliance. It was hypothesized that there is higher inter-rater reliability rating the alliance within CBT if raters have had work experience in the mental health field or a related practice and/or CBT training experience.

Method

Participants

Two months after Stage One of the study was conducted the same eight participants were recalled and split into two groups according to their clinical experience and CBT training. Two participants from each of these two groups were then randomly assigned to either Group C (i.e., 2x Clinical Experience and 2x No Clinical Experience) or Group D (i.e., 2x Clinical Experience and 2x No Clinical Experience). During the two month period between Stage One and Stage Two, participants 3 and 4 from Stage One of the study attended course paper Psychotherapy I and consequently obtained some CBT training experience before participating in Stage Two. However, both groups still appeared to be matched with each having a participant who had not experienced any CBT training or work experience. See Table 7.7 for the demographic characteristics of the participants in Group C and Group D in Stage Two of the experiment.

Table 7.7: Demographic Characteristics of the Participants in Group C and Group D
Group C

Participant	Age	Gender	Focus of post graduate study	CBT Training Experience	Work Experience
1	23yrs	Male	Clinical	Psychotherapy I	6 weeks
2	30 yrs	Female	Clinical	Psychotherapy I Various CBT workshops at NZPS conference	8 years
3	48yrs	Female	Clinical	None	None
4	47yrs	Male	Clinical	Psychotherapy I	None

Group D

Participant	Age	Gender	Focus of post graduate study	CBT Training Experience	Work Experience
5	52yrs	Female	Clinical	Psychotherapy I	2 years 6 months
6	24yrs	Female	Clinical	Psychotherapy I Padesky and Mooney's Workshop -resilience	2 years 4 months
7	50yrs	Female	Health	None	None
8	22yrs	Female	Clinical	Psychotherapy I	None

Measure

WAI-SR-O and Guidelines

The WAI-SR-O and the newly developed guidelines for rating the WAI-SR-O were also utilized for Stage Two of the experiment.

Experimental Design

When participants were recalled to take part in Stage Two of the study they were asked to not look at any previous notes from Stage One so that they all started on the same baseline. Participants in Group C were given the WAI-SR-O and the guidelines and participants in Group D were given the WAI-SR-O only. After a brief discussion all participants rated Session Four.

Statistical Analysis

Inter-rater Reliability

The use of Intraclass Correlation Coefficients (ICC; Haggard, 1958; Shrout & Fleiss, 1979) were continued for Stage Two. Models used included ICC (2, 8) for combined group calculations and ICC (2, 4) for individual group calculations. When ICCs were calculated for Clinical Experience and No Clinical Experience participants within each group ICC (2,2) was used.

Results

Table 7.8 presents ICCs for Group C and D and ICCs for participants with Clinical Experience and No Clinical Experience within those groups for Session Four. Overall Group D had a higher inter-rater reliability (0.60) than Group C (0.11). However, those participants with Clinical Experience within Groups C and D had higher inter-rater reliability (0.68) than those with No Clinical Experience (0.28).

Table 7.8: Intraclass Correlation Coefficients for Group C and D for Session Four

Session	Group C and D	Group C WAI-SR-O / Guidelines.	Group D WAI-SR-O only	Clinical Experience Group C and D	No Clinical Experience Group C and D
Session Four	0.63	0.11	0.60	0.68	0.28

Table 7.9 presents ICCs for Clinical Experience and No Clinical Experience participants within each group for Session Four. The ICC of participants with Clinical Experience in Group D was 0.83 and the ICC for those with Clinical Experience in Group C was 0.13.

Table 7.9: Intraclass Correlation Coefficients for Clinical Experience and No Clinical Experience participants within each group for Session Four

Session	Group C Clinical Experience	Group C No Clinical Experience	Group D Clinical Experience	Group D No Clinical Experience
Session Four	0.13	-0.57	0.83	0.06

Table 7.10 presents ICCs for Goal, Task and Bond subscales of the WAI-SR-O for Clinical Experience and Non Clinical Experience participants within each group for Session Four. Those participants with Clinical Experience within both Group C and Group D scored higher inter-rater reliability on the Goal subscale than participants with No Clinical Experience in Group C and Group D. Participants with Clinical Experience within Group D had higher ICCs on the Bond subscale. ICCs were 0.00 for the Task subscale of all groups indicating that within group variance equaled the between group variance.

Table 7.10: Intraclass Correlation Coefficients for Goal, Task and Bond subscales for Clinical Experience and Non Clinical Experience participants within each group for Session Four

Subscale	Group C Clinical Experience	Group C No Clinical Experience	Group D Clinical Experience	Group D No Clinical Experience
Goal	0.91	0.00	0.93	0.29
Task	0.00	0.00	0.00	0.00
Bond	0.11	-8.00	0.73	0.00

Discussion

This rater reliability study sought to investigate whether providing more information in the way of guidelines to rate the WAI-SR-O would lead to greater inter rater reliability. The alliance is a complex construct (Andrusyna, et al., 2001; Bordon, 1979; Gaston, 1990; Gelso & Carter, 1985, 1994; Luborsky, 1994) and rating this psychotherapy process can be very subjective (Horvath, 1994a; Horvath & Greenberg, 1994; Raue et al., 1993; Raue et al., 1991). Therefore, guidelines were developed to facilitate the rating of the WAI-SR-O within CBT. Stage One of the rater reliability study explored whether giving these clear, specific guidelines would increase inter-rater reliability. It was expected that training raters to use the newly developed guidelines to rate the WAI-SR-O would lead to higher inter–rater reliability for rating the alliance within CBT. However, the results did not support this hypothesis. Group A and Group B rated Session One utilizing the WAI-SR-O without the guidelines and Group B achieved higher inter-rater reliability than Group A. However, after this initial rating session overall ICC scores for both Group A and Group B at Session Two (i.e., after ICCs were recalculated for Group A with the scores of only three of the raters) and Session Three were similar, regardless of whether participants utilized the guidelines or not. Previous researchers (Bernardin et al., 1976; Blackburn et al., 2001; Nott et al., 1992; Raue et al., 1991) have found guidelines clarifying the rating of scale items beneficial and the findings of this rater reliability study have indicated that utilizing guidelines to rate the WAI-SR-O may account for some of the variability. However, in the current research, it was more likely that training rater's well (Athey & McIntyre, 1987; Bernadin & Pence, 1980; Kobak et al., 2008; Shohamy et al., 1992; Latham et al., 1975; Reichelt et al., 2003), giving them more opportunity to practice rating sessions (Hill & Stephany, 1990; Kim, 1995; Kobak et al., 2008) and sharing with them the expert's ratings and reasons for rating the items for the first session were more beneficial.

Rating error, particularly leniency error (McIntype, Smith, & Hassett, 1984; Saal & Landy, 1977), may have caused some of the raters to consistently score higher on the Bond subscale at Session One. The improvement of inter rater reliability for the Bond subscale at Sessions Two and Three may be due to feedback to the participants that their inter rater reliability was not as high on the Bond subscale. This improvement after feedback was in line with results of previous studies which suggested that feedback improved inter-rater reliability (Hill & Stephany, 1990; Kim, 1995). However, the improvement may also be due to sharing the expert's ratings and reasons for rating the items for the first session. Reliability for the Task subscale may have decreased as the participants concentrated on improving the reliability of the Bond subscale. However, these low or negative ICCs for the task subscale may also have indicated a need for additional training and/or further improvements to the guidelines.

The ICC's for both Group A and Group B decreased at Session Three. This could be interpreted as participants in both groups suffering from information overload and/or fatigue at the time of this rating. The lowering of the ICCs for the Goal, Task and Bond subscales for Group B at Session Three may have been due to these participants having learnt to rely on their clinical judgment at Sessions One and Two and then being overwhelmed and/or distracted by the specifics of the guidelines at Session Three. Most participants reported that it was easier to rate the WAI-SR-O with the guidelines. However, some of the participants pointed out that after being trained with the guidelines, they found learning what the specific items meant; learning how to rate the items and then rating a session at the same time was difficult. This factor may have handicapped Group A in Session Two, whereas by Session Three they had gained more confidence and with more practice inter-rater reliability continued to improve for the Goal and Bond subscales. In sum, while some positive feedback was provided by the participants about the utilization of the guidelines, the guidelines did not appear to significantly increase inter-rater reliability. This finding is similar to that of raters utilizing the CTS-R (Blackburn et al., 2001) to evaluate CBT trainees in the Newcastle Cognitive and Behavioural Therapies Centre and other Cognitive and Behavioural Therapies centres throughout the United Kingdom. On one hand raters found the more detailed CTS-R more helpful, but on the other hand they found there was more room for discrepancy (P. Armstrong, personal communication, 15th October, 2009).

After examining the demographic differences between both groups it appeared that Group B had more participants with work experience and CBT training experience. Previous studies have demonstrated that the characteristics of raters can affect inter-rater reliability (Feldman, 1978; Hambleton & Powell, 1983; Kiesler, 1973; Kim, 1995; McIntype et al., 1984; Moras & Hill, 1991; Moritsch & Suter, 1988; von Raffler-Engel, 1980). In particular, Kim (1995) found that selecting raters who were knowledgeable about the topic being rated committed fewer rating errors. In Stage One ICC calculations provided evidence that participants who have had some work experience and/or CBT training experience were more reliable to rate the alliance with or without the guidelines. This was explored further in Stage Two.

The aim of Stage Two was to investigate whether there was higher inter-rater reliability rating the alliance within CBT if raters have had work experience in the mental health field or a related practice and/or CBT training experience. The current research was limited by the small number of participants, particularly when calculating the ICCs for Clinical Experience and No Clinical Experience for each of the groups. Nervertheless, the findings demonstrated that participants with Clinical Experience had higher inter-rater reliability than those with No Clinical Experience. However, the inter-rater reliability of the participants with Clinical Experience in Group D was higher than that of participants with Clinical Experience in Group C, with the main difference being in the rating of the Bond subscale. It is possible that other individual character differences, such as specific academic background, specific work experience, life experience, personality, memory retention, and/or learning style may also play a part in determining rater error.

To date most psychotherapy process-outcome research has relied on early research from applied psychology conducted within organizational and educational settings to found the basis of their rater selection and training. Therefore, it has been difficult to determine whether the results of this early research is relevant to selecting and training raters to rate processes within psychotherapy. The findings of this rater reliability study recall the review of rater selection in psychotherapy process research of Moras and Hill (1991) who found that the level of experience needed by raters was directly related to the amount of inference required by the task. Furthermore, the utilization of clinically experienced raters within counselling research studies (Burstein & Carkhuff, 1968; Carkhuff et al., 196; McMullin, 1972; Mitchell & Brenson, 1970) has also suggested that clinical training and experience may be required to rate the more abstract constructs, such as alliance.

Results could indicate that the WAI-SR-O may not adequately assess alliance processes within CBT, particularly the Task component of the alliance. The WAI is now over twenty five years old and CBT and the role of the alliance within CBT has evolved immensely over this time. Bordon (1979) also indicated that particular therapies such as CBT may impose unique demands on therapist-client relationship and that alliance measures may fail to accurately assess the alliance at more differentiated levels. However, the WAI-SR-O was developed for the current research and there is no known exploratory factor analysis to gain better understanding of its factor structure. Future research could also explore the suitability of other alliance scales or consider developing a new scale to capture alliance processes that align with the collaborative empiricism of CBT (Beck et al., 1979).

In conclusion, although the sample size of this rater reliability study and the possibility that the newly developed guidelines may not be appropriate for rating the WAI-SR-O within other psychotherapies limits the generalisability of the findings it does offer further understanding toward selecting and training raters, reducing rater error and increasing the inter-rater reliability for rating the alliance within CBT. Guidelines for rating the WAI-SR-O did not significantly increase inter-rater reliability and improving rater reliability was more reliant upon training raters well, providing the opportunity to rate practice sessions and apprenticeship for raters in training. The rater reliability study also demonstrated the need to select raters with some clinical experience when rating the alliance, but future process-outcome research may need to address specific character differences for reducing error and increasing inter-rater reliability. Future researchers could further improve the guidelines for rating the WAI-SR-O, particularly the Task subscale, conduct a factor analysis of the WAI-SR-O to gain a better understanding of its factor structure, or alternatively consider the suitability of other alliance scales (e.g., CALPAS) or the development of a new alliance scale to capture the alliance construct within CBT.

Chapter 8 reviews the inter-rater reliability for the WAI-SR-O in the current research and presents findings across the case studies.

Chapter 8: Results of the Current Research

Overview

Chapter 8 is divided into two sections. The first section reviews the inter-rater reliability for the WAI-SR-O in the current research. The second section sets out case results. Graphed sessions of depression severity level, Total Alliance, and the Goal, Task and Bond subscales of the WAI-SR-O are presented. Trends are described in the following sequence: depression severity scores, Total Alliance scores and Goal, Task and Bond subscale scores, depression severity and Total Alliance, and depression severity and Goal, Task and Bond subscales. Individual graphs are presented, rather than graphs of averaged scores which tend to have the effect of smoothing out variability.

Background information for each case is outlined more fully [2] and individual graphed sessions of depression severity, Total Alliance and Goal, Task and Bond subscales and their trends are presented in Appendix F.

---

[2] Names and any identifying information have been changed to protect the client's identity. Please note the words 'participant' and 'client' can be used interchangeably in the results and discussion sections of this thesis. However, to simplify this issue and differentiate the 10 participants in the current research from the 8 participants in the rater reliability study, the 10 participants in the current research will now be referred to as clients.

Preliminary Analyses: Inter-rater reliability

Intraclass Correlations Coefficients for the WAI-SR-O

The current research relied upon observer ratings and therefore it had to be demonstrated that the raters were able to identify whether targeted therapist-client behaviours occurred in the viewed sessions. In order for there to be confidence in analyses stemming from ratings made by the use of independent raters there needed to be a high level of agreement between the raters (Kaplan & Saccuzzo, 2001). Three recalibrations were conducted for the WAI-SR-O in the current study to facilitate rater retraining.

Table 8.1 presents a summary of the ICCs for overall inter-rater reliability and inter-rater reliability of the Goal, Task and Bond subscale scores during the recalibration of 10, 20, and 30 consecutive ratings. The overall reliability of the recalibration of the first set of ten ratings indicated a satisfactory level of agreement with an ICC of 0.76. The inter-rater reliability of the Goal and Bond subscales also indicated a satisfactory level of agreement with ICCs of 0.78 and 0.60 respectively. The inter-rater reliability of the Task subscale indicated an acceptable, but possibly improvable level of agreement with an ICC of 0.49. However, the overall inter-rater reliability and the inter-rater reliability of the Goal, Task and Bond subscales decreased in the second set of ten ratings indicating an acceptable, but possibly improvable of agreement. After feedback and some retraining, the overall inter-rater reliability rose to a satisfactory level of agreement again of 0.69 in the third set of ten ratings. The inter-rater reliability Goal, Task and Bond subscales also indicated a satisfactory level of agreement with ICCs of 0.69, 0.71, and 0.63 respectively (Landis & Koch, 1977; Portney &Watkins, 1993; Rosner, 2006).

Table 8.1: Intraclass Correlation Coefficients for the Recalibration of 10, 20 and 30 Consecutive Ratings

	Overall Reliability	Goal	Task	Bond
Recalibration for 1st set of ten ratings	0.76	0.78	0.49	0.60
Recalibration for 2nd set of ten ratings	0.49	0.48	0.47	0.52
Recalibration for 3rd set of ten ratings	0.69	0.69	0.71	0.63

Table 8.2 summarizes ICCs during recalibration of items for 10, 20, and 30 consecutive ratings of the WAI-SR-O. Recalibration for 1st set of ten ratings indicated that further training was needed for item 4 of the Goal subscale, items 5 and 7 of the Task subscale and item 12 of the bond subscale. Recalibration for the 2nd set of ten ratings indicated that further training was needed for most items. However, raters reported that during some of these sessions they had difficulty hearing therapist and client interactions due to noise recorded onto the DVD. Recalibration for 3rd set of ten ratings indicated that further training was still needed for item 4 of the Goal and item 5 of the Task subscale. They were also having difficulty with items 9 and 11 of the Bond subscale.

Table: 8.2 Intraclass Correlation Coefficients for the recalibration of items for 10, 20 and 30 consecutive ratings

Items	Recalibration for 1st set of ten ratings	Recalibration for 2nd set of ten ratings	Recalibration for 3rd set of ten ratings
1	0.90	0.44	0.61
2	0.79	0.54	0.73
3	0.66	-0.42	0.64
4	0.27	0.51	0.47
5	0.38	0.83	0.25
6	0.55	-0.42	0.84
7	0.36	-0.12	0.54
8	0.61	0.69	0.75
9	0.66	0.44	0.42
10	0.56	0.55	0.67
11	0.82	0.60	0.47
12	0.00	0.44	0.81

Table 8.3 presents a summary of the Intraclass Correlations Coefficients for the WAI-SR-O in the current research. The overall inter-rater reliability indicated a satisfactory level of agreement (Landis & Koch, 1977; Portney &Watkins, 1993; Rosner, 2006) with an ICC of 0.68. The inter-rater reliability Goal and Bond subscales also indicated a satisfactory level of agreement with ICCs of 0.74 and 0.64 respectively. The inter-rater reliability of the Task subscale indicated an acceptable, but possibly improvable level of agreement with an ICC of 0.54.

Table 8.3: Summary of Intraclass Correlations Coefficients for the WAI-SR-O for the current research

Overall Reliability	Goal	Task	Bond
0.68	0.74	0.54	0.64

Overall, the inter-reliability of the Task subscale for the current research was problematic, which is interesting considering Hatcher and Gillaspy (2006) reported that the Task subscale of the WAI-SR better differentiated the task dimension. The lowered inter-rater reliability of the Task subscale maybe due to one of or a combination of the following: a) not training the raters adequately, b) the retraining of raters after the 2nd recalibration not being adequate and/or sustainable, c) noise recorded onto the DVD, d) rater fatigue after rating three sessions back to back for the first 20-30 sessions, e) only one of the raters having some clinical experience and f) the WAI-SR-O and/or guidelines not effectively assessing the Task dimension of the alliance in the current research. Training guidelines and how the instrument is implemented is a crucial variable affecting the utility of the instrument (L. R. Fenton, personal communication, 23rd August, 2010). The definition of the task dimension (i.e.,"supportive/reassurance, reflection, reformulation") from the original guidelines (i.e., page one of the guidelines) of Raue and colleagues (1991) (see also Raue & Goldfried, 1994) has been criticized as being confused with the definition of the bond dimension from Bordin's perspective (L. R. Fenton, personal communication, 23rd August, 2010) and therefore the raters of the current research may not have been adequately trained. However, it is possible that these tasks are reflective of the tasks of various other psychotherapy orientations. Furthermore, the raters of the current research were thoroughly trained to observe the task dimension of the alliance within CBT. Instances of the Task subscale were demonstrated from recorded CBT sessions from the Depression Outcome Study and then related back to these items of the WAI-SR-O. Moreover, a low ICC does not necessarily indicate poor overall agreement, but may reflect the use of the absolute agreement option which has more stringent criteria than the estimated rater agreement option (Hall, Groome, Streiner, & Rochon, 2006). When the raw data was eyeballed, it appeared that if the raters were not agreeing on a score they were only one point apart. There were only 2-3 scores in which they were 2 points apart. The lowering of the overall task subscale inter-rater reliability for the current research will be discussed further in the discussion section.

Consistencies Across Cases

Total Alliance scores of the WAI-SR-O range from 12 to 60. While these scores have not been broken down into differential ranges by past researchers, the current research has utilized the ranges of 12 to 24 (low range, or seldom/sometimes), 25 to 36 (moderate range or fairly often), 37.5 to 48 (High range or very often) and 48.5 to 60 (Very High range or always) to facilitate the detailing of results. This may also facilitate showing whether the strength (i.e. range of the WAI-SR-O) of the alliance has a role in predicting outcome.

Scores for each subscale range from 1-20. To facilitate detailing trends of the Goal, Task and Bond subscale scores of the WAI-SR-O the current research has utilized the following ranges of 1-4 (Very Low range or seldom), 4.5-8 (Low range or sometimes), 8.5-12 (Moderate range or fairly often), 12.5-16 (High range or very often) and 16.5-20 (Very High range or always).

As outlined earlier depression severity was rated by the client at the beginning of each session. Scores described the clients mood, cognition and behaviour over the past week including the day of the session. Alliance or the interaction between the therapist and the client was rated by the raters utilizing the WAI-SR-O in that session.

Depression Severity Symptom Change
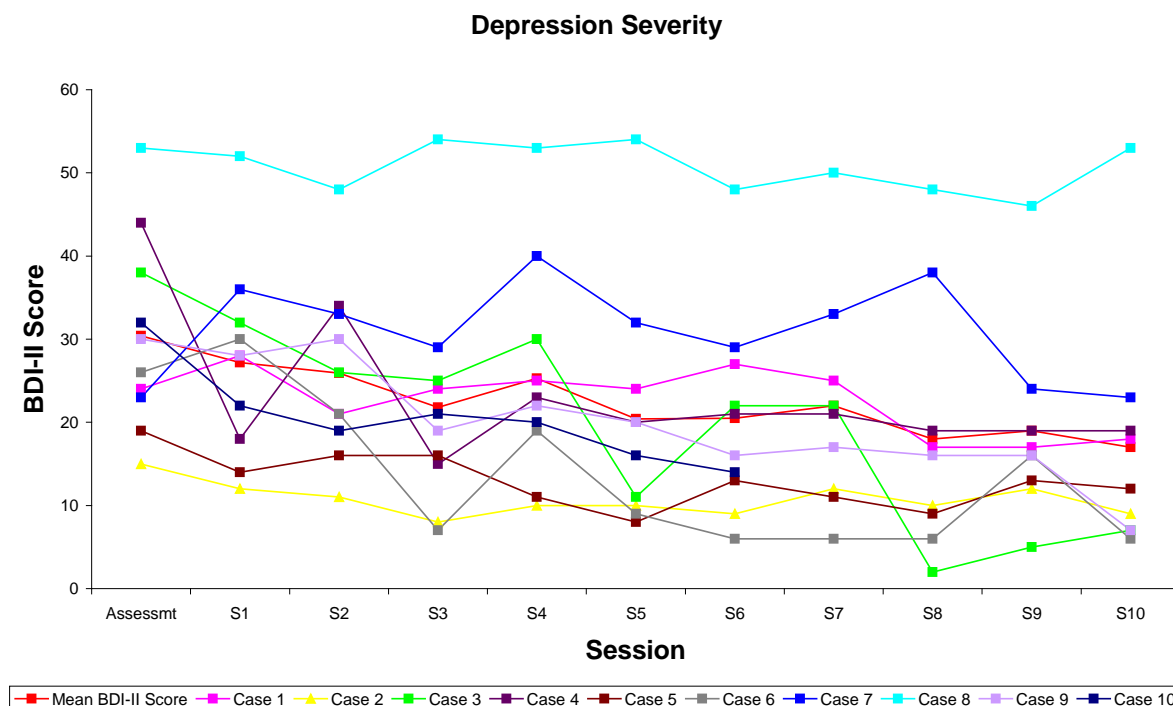
Figure 8.1: Individual depression severity patterns

Figure 8.1 graphically presents the depression severity scores of the ten clients. At initial clinical assessment all clients were diagnosed with major depression with depression severity ranging from mild to severe scores on the BDI-II. There were two cases with mild scores, three cases with moderate scores and five cases with severe scores. Eight clients (i.e., cases 1, 3, 4, 6, 7, 8, 9, 10) were assessed on the CIDI as also having a comorbid anxiety disorder. The two cases (i.e., cases 2, 5) with mild scores on the BDI-II were not assessed as having a comorbid anxiety disorder. The depression severity of the three cases (i.e., cases 1, 6, 7) within the moderate range on the BDI-II had one or two comorbid anxiety disorder(s), usually panic disorder without agoraphobia and/or generalized anxiety disorder. The depression severity of the three cases (i.e., cases 4, 8, 9) within the severe range on the BDI-II had two or three comorbid anxiety disorders including: panic disorder without agoraphobia, generalized anxiety disorder, obsessive-compulsive disorder and post traumatic stress disorder. The client in case 3 experienced severe depression, GAD, PTSD and bulimia nervosa. The client in case 10 experienced severe depression and generalized anxiety disorder. Nine clients finished therapy with either no depression or mild depression according to their BDI-II scores. The client in case 8 finished therapy with her depression

severity in the lower end of the severe range. Table 8.4 presents a summary of depression severity at initial assessment and therapy completion.

Table 8.4: Summary of depression severity at assessment, Session 10 and completion of therapy

Client	Depression Severity at Pre-assessment	Depression Severity at Session 10	Depression Severity at the End of Therapy
1	Moderate	Mild	Mild
2	Mild	No Depression	No Depression
3	Severe	No Depression	No Depression
4	Severe	Mild	Mild
5	Mild	No Depression	No Depression
6	Moderate	No Depression	No Depression
7	Moderate	Moderate	No Depression
8	Severe	Severe	Lower end of Severe
9	Severe	No Depression	No Depression
10	Severe	-	Mild

In seven cases (i.e., cases 2, 3, 4, 5, 8, 9, 10) depression severity decreased between pretreatment assessment and the first therapy session. In three cases (i.e., cases 1, 6, 7) depression severity increased between pretreatment assessment and the first session.

The patterns of depression severity of the clients across the ten sessions (six sessions in case 10) were mixed (see Figure 8.1). While there was a gradual decline in depression severity in nine of the cases, the pattern of change was discontinuous and non-linear. Over the first week of therapy depression severity decreased in seven cases (i.e., cases 1, 2, 3, 6, 7, 8, 10) and increased in three cases (i.e., cases 4, 5, 9). After an initial decrease, the depression severity of three clients (i.e., cases 3, 6, 10) continued to gradually decrease with one or two peaks to session 10. The depression severity of three other clients (i.e., cases 2, 7, 8) gradually decreased but scores on the BDI-II either fluctuated each session or fluctuated weekly to session 10. The depression severity of the client in case 1 fluctuated over the first half of the sessions and then decreased in the last four sessions.

For two clients (i.e., cases 4, 9) who experienced an increase of depression severity over the first week of therapy there was a gradual decrease of depression severity with one or two peaks of depression severity for the remainder of the ten sessions. The depression severity of the client in case 5 decreased gradually over the first half of the sessions and then fluctuated weekly in a linear manner over the last five sessions.

Total Alliance
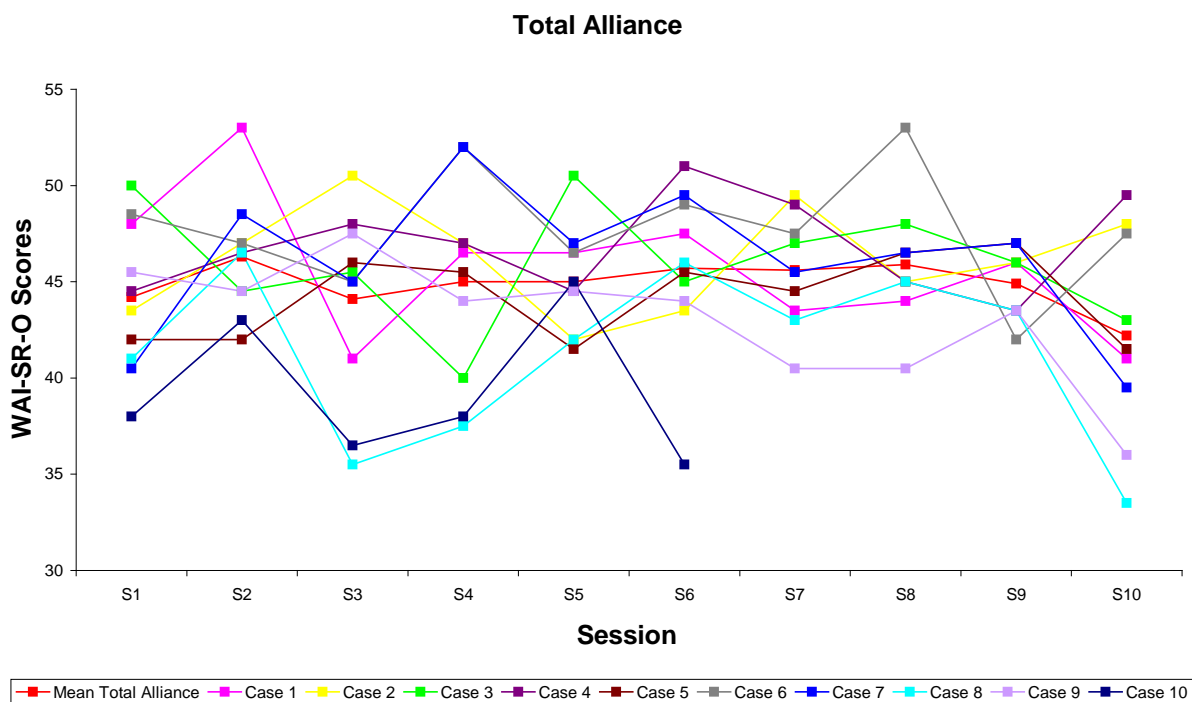
Figure: 8.2 Individual Total Alliance patterns

Figure 8.2 graphically presents Total Alliance scores for the ten clients. Appendix G presents tables of Total Alliance and subscale scores to facilitate viewing of the results. Overall, Total Alliance scores on the WAI-SR-O fell between the High range and the lower end of the Very High range. The Total Alliance score for a majority of the research sample (i.e., cases 1, 2, 3, 4, 6, 7) fell in the High range of the WAI-SR-O throughout the 10 sessions, with some of the scores falling into the lower end of the Very High range. In case 5 all the Total Alliance scores fell into the High range. In three cases (i.e., cases 8, 9, 10) the Total Alliance scores fluctuated between the higher end of the Moderate range and the High range. However, the Total Alliance score fell into the higher end of the Moderate range for only four sessions out of the 26 sessions rated for these three cases.

Total Alliance scores in the initial therapy session mostly fell in the High range, but the Total Alliance scores for cases 3 and 6 fell into the lower end of the Very High range during this session. In the first week of therapy the Total Alliance score in six cases (i.e., cases 1, 2, 4, 7, 8, 10) increased, in one case (i.e., case 5) the score was maintained and in the other three cases (i.e., cases 3, 6, 9), the score decreased, but the Total Alliance score remained in the High range. For a majority of the cases the Total Alliance score appeared to peak by the third session. For the remainder of the sessions the Total Alliance scores for all ten cases either fluctuated weekly or each session in a fairly stable manner.

Total Alliance and Depression Severity

There appeared to be four different alliance-depression severity patterns that fluctuated in no particular order throughout the 10 sessions (6 sessions in case 10). All of these patterns occurred in the beginning, middle and end of the 10 sessions. The random manner in which the alliance-depression-severity patterns occurred throughout the therapy sessions was different for each case.

In all ten cases there were sessions in which there was a decrease in depression severity and an increase in Total Alliance score ($\downarrow$ depression severity, $\uparrow$ Total Alliance) in the same session. The Total Alliance score usually fell in the High range or lower end of the Very High range in the preceding session and in the same session as the rated depression severity. This pattern first occurred in either the second or third session for a majority of the sample (e.g., cases 1, 2, 3, 4, 7, 8, 9, 10). While this thesis does not have the scope to investigate the content of therapy sessions alongside alliance and depression severity data it will attempt to give some case examples of session content which may contribute toward alliance-depression severity patterns. In the first session of case 1 the client received psychoeducation surrounding CBT, the 5 part-model was demonstrated and and a problem list and goals for her therapy were explored. The Total Alliance score fell in the High range during this session. For homework the client continued to think about her therapy goals. The client rated her depression severity lower at the beginning of the next session (session 2). In session 2 the client and therapist examined a very embarrassing situation for the client utilizing the 5 part-model intervention. The Total Alliance score increased into the Very High range for this session. Between the 7th session (Total Alliance score fell in the High range in session 7) and 8th session of case 1 the client used skills she learnt in an earlier session in therapy to resolve a workplace conflict. Her depression severity was

lower at the beginning of session 8 and the alliance increased in that session. Furthermore, between the 4th session (Total Alliance score fell in the High range in session 4) and 5th session of case 3 the client went on a pleasure trip for a week which she enjoyed immensely. Her depression severity dropped 19 points on the BDI-II at the beginning of session 5 and the alliance increased into the Very High range.

In nine cases (i.e., cases 1, 2, 3, 4, 6, 7, 8, 9, 10) there were sessions in which there was an increase in depression severity and a decrease in Total Alliance score (↑ depression severity, ↓Total Alliance) in the same session. When this pattern occurred there was usually a gradual building of the Total Alliance scores over 2-3 sessions preceding the increase in depression severity. The Total Alliance score mainly fell in the High range or lower end of the Very High range in the preceding session and in the same session as the rated depression severity. Mixed alliance–depression severity patterns occurred in the following session. In a majority of the cases there was a decrease in depression severity and an increase in Total Alliance scores. However, in other cases the depression severity increased and the Total Alliance scores decreased, or both the depression severity and Total Alliance scores increased, or both the depression severity and Total Alliance scores decreased. For example, in case 1 the client's depression severity had decreased from the previous session and the alliance was in the Very High range in session 2. However, in session 3, the client rated her depression higher and the alliance dropped to the lower end of the High range. Between sessions 2 and 3 the client had had relationship problems with her partner and other family members. She was also concerned about her own angry outbursts. The therapist and client worked with these issues during session 3. The client's depression severity continued to increase in session 4, but the Total Alliance score increased. Between sessions 3 and 4 the client reported that she had had to deal with a racist comment from a work colleague and that she was still felt angry about it. For homework the client had been working on her activity schedule. In session 4 the client and the therapist discussed the incident at work and reviewed the client's activity schedule. They discovered that she became angrier in the afternoons, but that if she knew she was going to the gym after work she felt better.

In all ten of the cases there were sessions in which there was a decrease in depression severity and a decrease in Total Alliance score (↓ depression severity, ↓ Total Alliance) in the same session. The Total Alliance score mostly fell in the upper end of the High range or the lower end of the Very High range in the preceding session and in the same session as

the rated depression severity. In the following session the Total Alliance score usually increased and the depression severity decreased. However, sometimes the depression severity increased (e.g., cases 4, 8, 9, 10) and/or the Total Alliance scores decreased (e.g., cases 8, 9, 10). For example, in case 1 the Total Alliance score in session 6 fell in the High range. The client rated her depressive symptoms lower at the beginning of session 7, but the Total Alliance score decreased in that session. In session 7 the client was concerned about not completing her homework and what the therapist would think about her for not doing her homework. Furthermore, thought records had been introduced in the session as an intervention and there had not enough time to practice a 3 column thought record so the client was not clear about what she had to do for homework between the two sessions. These issues may have created ruptures in the alliance. Resolving these ruptures within session 7 and/or the fact that the client utilized her skills to resolve her workplace conflict between sessions 7 and 8 and thereby gaining a sense of mastery, may have contributed to a continued decrease in depression severity and an increase alliance in the following session.

Finally, in seven of the cases (i.e., cases 1, 2, 4, 5, 6, 7, 8) there were sessions in which there was an increase in depression severity and an increase in Total Alliance score (↑ depression severity, ↑ Total Alliance) in the same session. The Total Alliance score mostly fell into the upper end of the High range or the lower end of the Very High range in the preceding session and in the same session as the rated depression severity. Mixed alliance–depression severity patterns occurred in the following session including: decrease in both depression severity and Total Alliance scores, increase in both the depression severity and Total Alliance scores, decrease in depression severity and an increase in Total Alliance scores and an increase in depression severity and a decrease in Total Alliance scores. For example, in case 4 the client's depression severity increased 16 points on the BDI-II in session 2. The client had had relationship problems with her husband between session 1 and 2. The therapist and client discussed these problems in session 2 and the Total Alliance increased in session 2 and continued to increase in session 3. Her depression severity dropped 19 points on the BDI-II at the beginning of session 3.

Overall Summary

There were four main trends that fluctuated in no particular order throughout the 10 sessions (6 sessions in case 10), but there was little consistency of alliance–depression severity patterns within the cases.

Goal Subscale

Goal Component of the Alliance
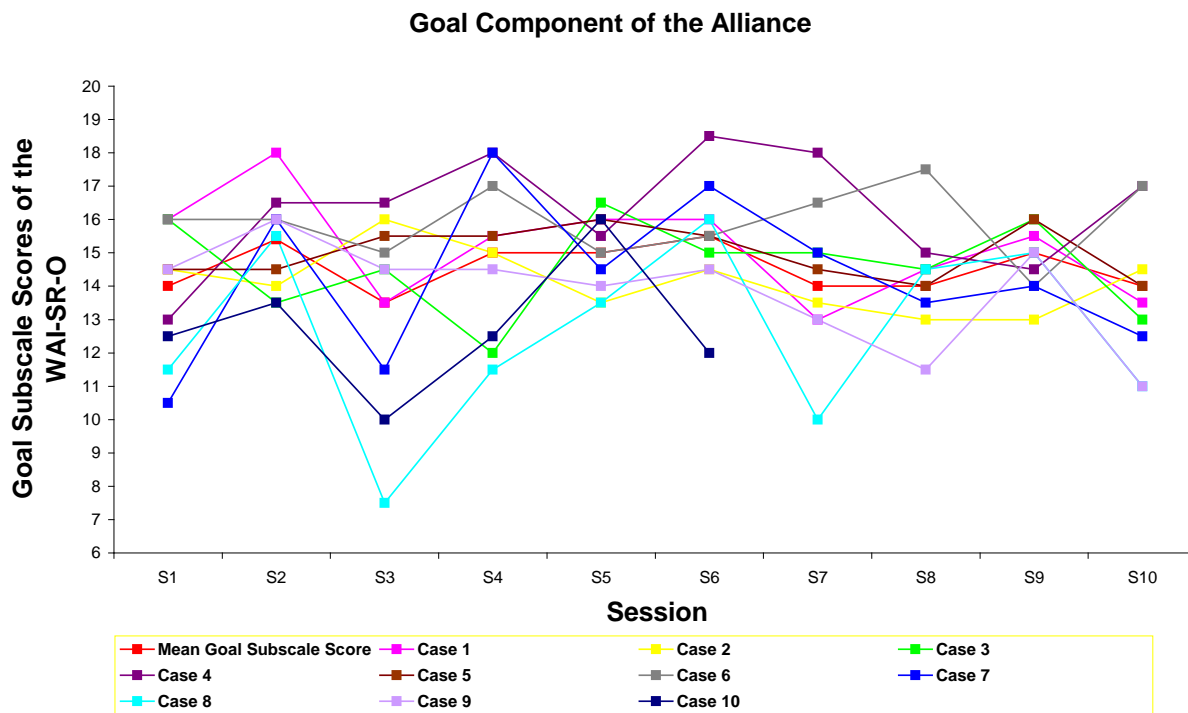
Figure: 8.3 Individual Patterns of the Goal Component of the Alliance

Figure 8.3 graphically presents Goal subscale scores for the ten cases. Overall the scores of the Goal subscale fell between the Moderate range and the lower end of the Very High range. In three cases (i.e., cases 1, 2, 5) the scores of the Goal subscale fell mostly in the High range with some scores in the Very High range. In two cases (i.e., cases 4, 6) the scores fell equally between the High range and the Very High range. In four cases ((i.e., cases 3, 7, 9, 10) the scores fell mostly in the High range with some scores in the Moderate range. In case study 8 the scores fell equally in the High range and Moderate ranges, with one score falling into the Low range.

Task Subscale
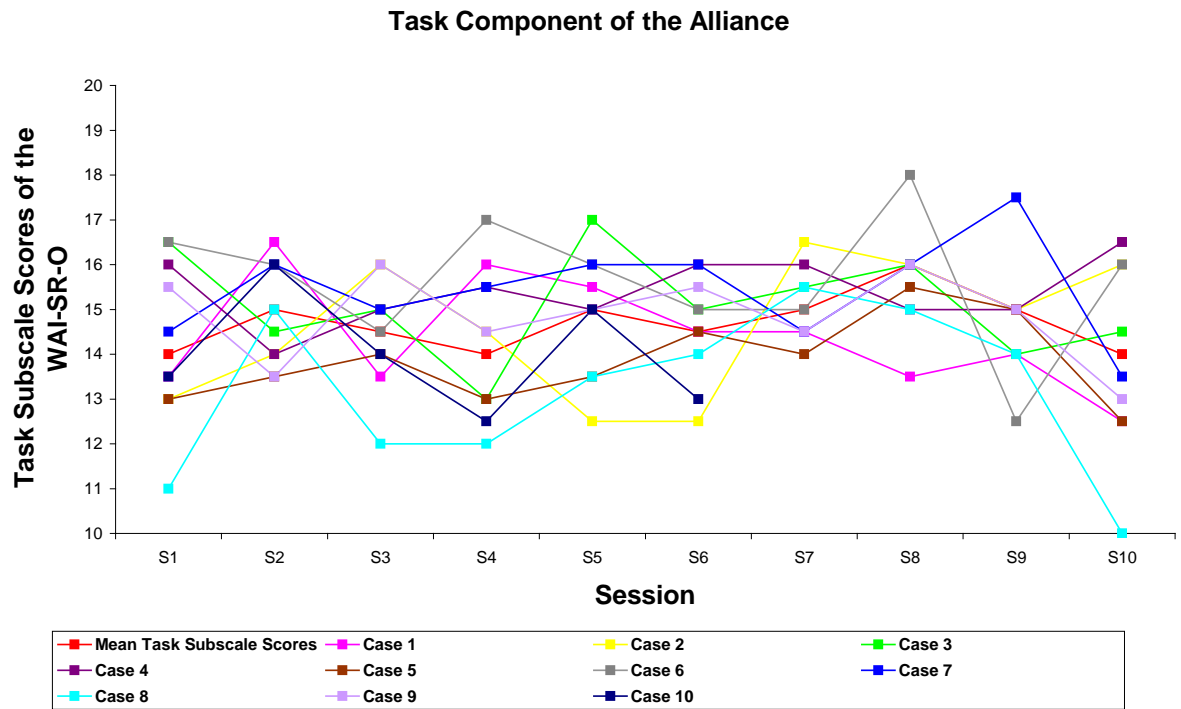
Figure: 8.4 Individual Patterns of the Task Component of the Alliance

Figure 8.4 graphically presents Task subscale scores for the ten cases. Overall the scores of the Task component of the Alliance fell into the High range with a few scores falling into the Moderate range and lower end of the Very High range. In nine of the cases (i.e., cases 1, 2, 3, 4, 5, 6, 7, 9, 10) the scores of the Task subscale fell mostly in the High range with one or two scores falling into the lower end of the Very High range. Case 6 had the most scores in the lower end of the Very High range. In case 8 the scores fell mostly in the High range, but some scores fell into the Moderate range. Notably, there is little variability in the scores. This lack of variance will be discussed further in the Limitations of the Current Research and Future Research Recommendations section of the discussion.

Bond Subscale

Bond Component of the Alliance
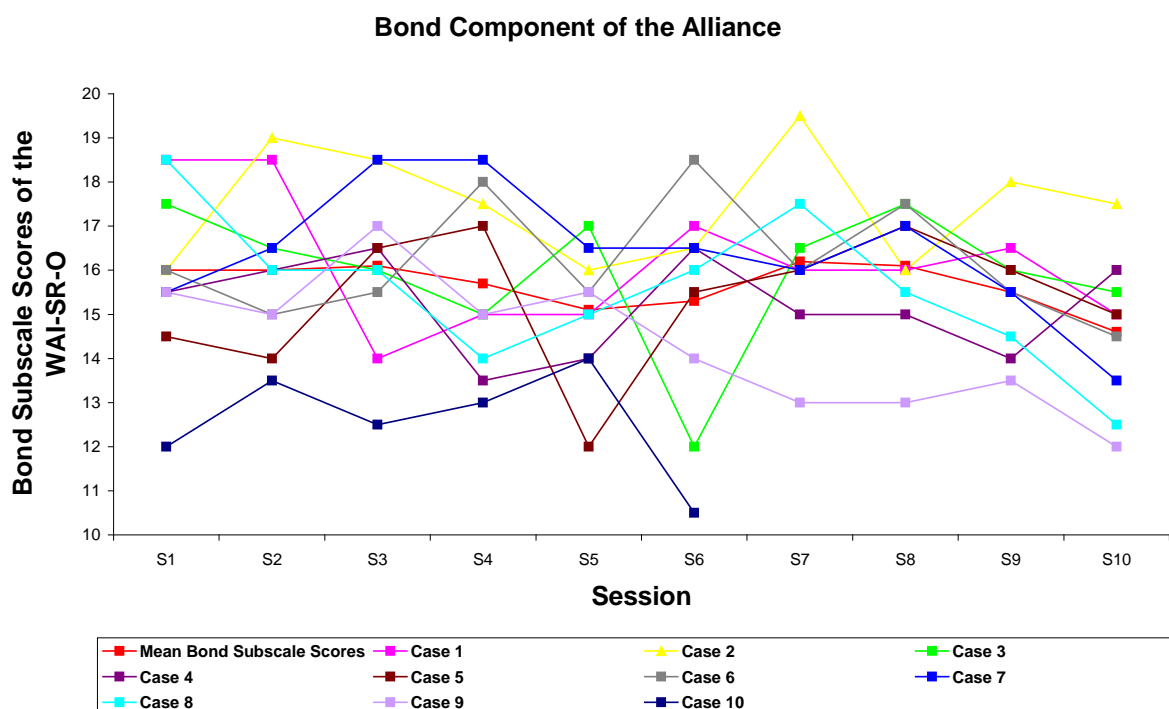
Figure 8.5: Individual Patterns of the Bond Component of the Alliance

Figure 8.5 graphically presents Bond subscale scores for the ten cases. Overall the scores of the Bond component of the Alliance fell between the upper end of the Moderate range and lower end of the Very High range. In a majority of the cases (i.e., cases 1, 4, 5, 6, 8, 9) the scores of the Bond subscale fell mostly in the High range. In three cases (i.e., cases 2, 3, 7) the scores mostly fell in the Very High range with some scores falling into the High range. In case study 10 the scores fell mostly in the High range with a couple of scores falling into the Moderate range.

Development of Goal, Task and Bond Components

Overall the Bond component of the alliance appeared to be the most developed (i.e., higher subscale scores) throughout the ten sessions for a majority of the case studies (i.e., cases 1, 2, 3, 5, 7, 8). If the quality of the Bond component of the alliance is low interactions between the client and the therapist may initially be built through being high on the Goals and/or Tasks (Pinsoff, 1994). For case 9 the Bond component of the alliance appeared to develop in the first five sessions and then in the last five sessions the Task component developed more. The Task component of the Alliance appeared to be the most developed in case 10. The Goal component of the Alliance appeared to be the most developed in case 4. See Figure 8.6 for graphical display of the Goal, Task and Bond subscale scores for each case.

Goal Items

In all the cases there were sharp and/or gradual increases and decreases of Goal subscale scores. However, when raw data was eyeballed, there was no clear pattern of which items (i.e., client and therapist collaboratively setting goals for the session, agreement on what was important for the client to work on, client and therapist working on mutually agreed, client and therapist have established a good understanding of changes that would be good for the client) were influencing these increases or decreases. Sometimes it appeared that all four goal items may have been involved in the decrease or increase in the Goal subscale score and at other times one, two or three items may have been involved. Items of the Goal subscale and the number of items involved in the decrease or increase of the subscale were different for each case and different from session to session.

Task Items

In all the cases there were sharp and/or gradual increases and decreases of Task subscale scores. When raw data was eyeballed, there was no clear pattern of which items (i.e., agreement on the usefulness of the current activity in therapy, agreement that what the client and therapist are doing in therapy will help the client to accomplish the changes he/she needs, as a result of the sessions there is clarity about how the client might be able to change, the client believes that the way they are working on his/her problem is correct) were influencing these increases or decreases. Sometimes it appeared that all four goal items may have been involved in the decrease or increase in the Task subscale score and at other times one, two or three items may have been involved. Items of the Task subscale

and the number of items involved in the decrease or increase of the subscale were different for each case and different from session to session.

Bond Items

In all the cases there were sharp and/or gradual increases and decreases of Bond subscale scores. When raw data was eyeballed in a majority of the cases, two items (i.e., client feels the therapist respects and cares about the client even when the client does things that the therapist does not approve of, client feels that the therapist appreciates him/her as a person) appeared to be involved in a sharp decrease in Bond subscale scores. However, these two items and items reflecting a mutual liking and respect between the client and the therapist were usually involved in a sharp increase in the Bond subscale score.

Overall Summary

The Bond component of the alliance appeared to be the most developed throughout the ten sessions for a majority of the cases (i.e., cases 1, 2, 3, 5, 7, 8). However, there was little consistency around which items were involved in increases and decreases of the Goal, Task and Bond subscale scores.

Case 1
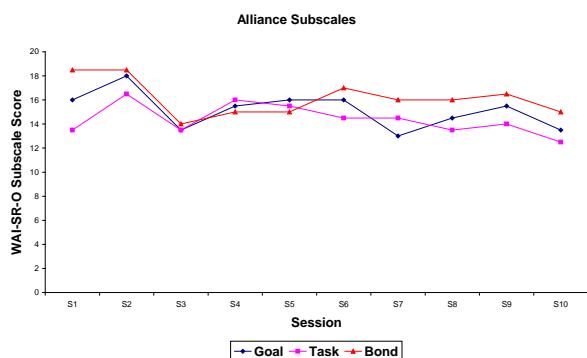
Case 2
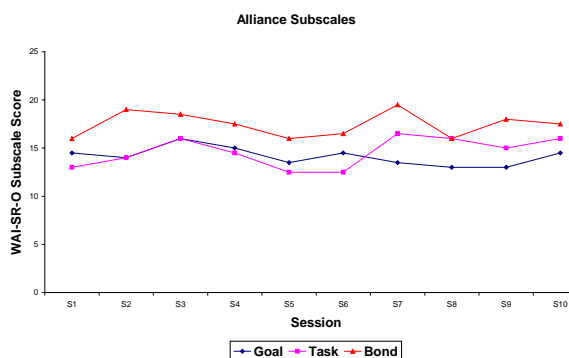
Case 3
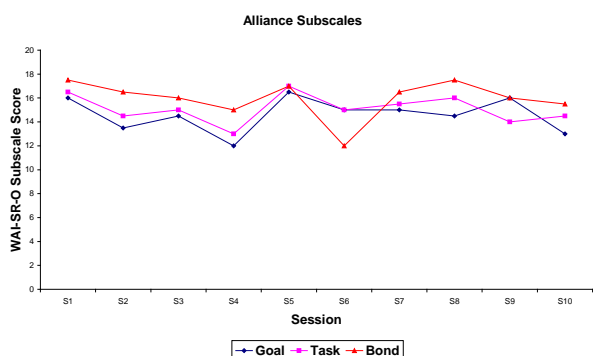
Case 4
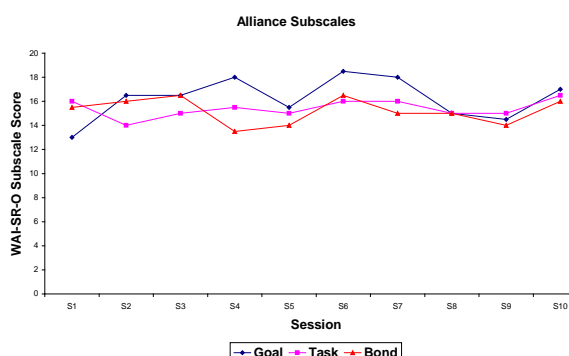
Case 5
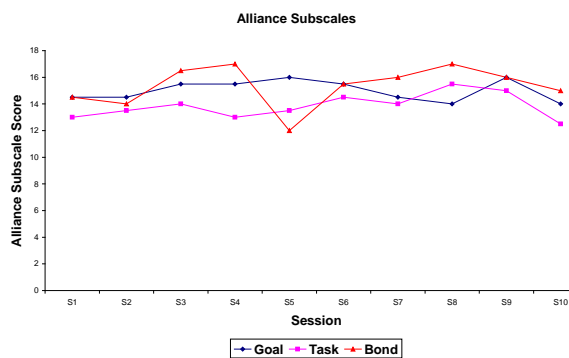
Case 6
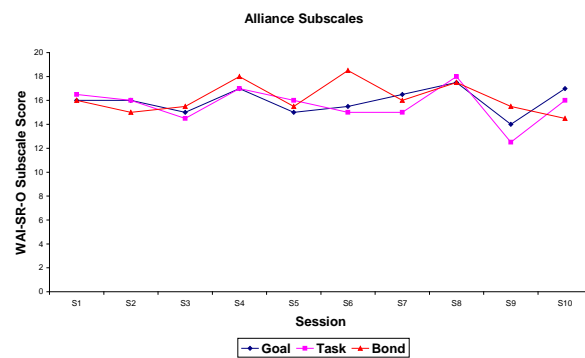
Case 7
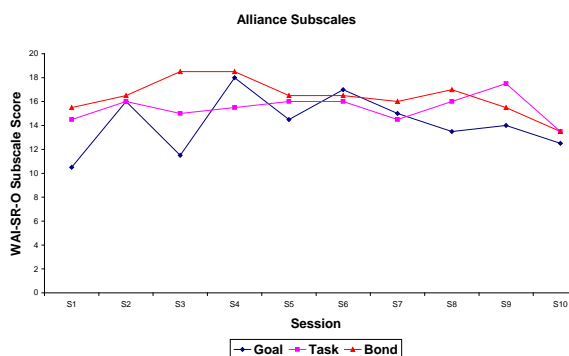
Case 8
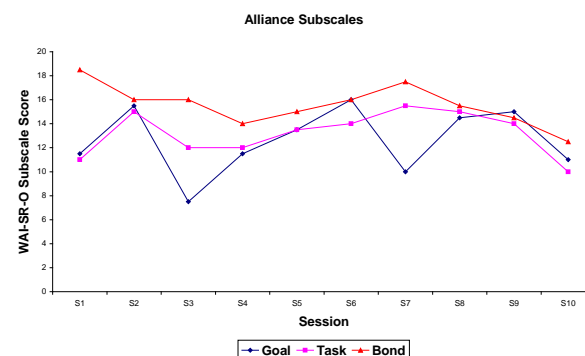
Case 9
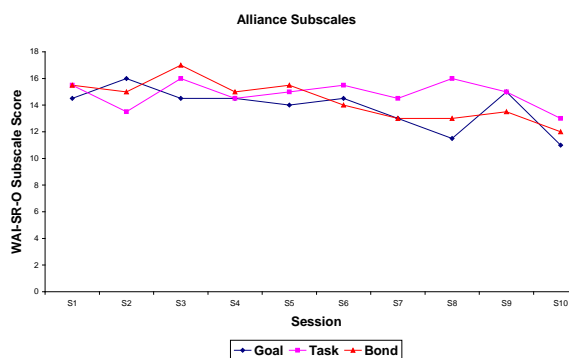
Case 10
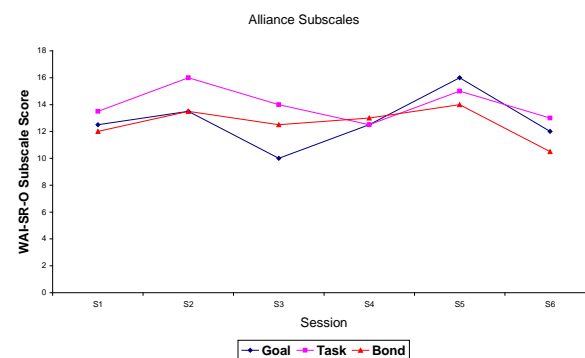
Figure: 8.6 Graphical Display of the Goal, Task and Bond Subscale Scores for Each Case

Goal, Task, and Bond Subscales and Depression Severity

There appeared to be five main Goal, Task, and Bond subscales-depression severity patterns that fluctuated throughout the 10 sessions (6 sessions in case 10). The random manner in which these patterns occurred was different for each case. Furthermore, these patterns were not clearly defined and it was difficult to determine typical trends for the following session.

In nine cases there were sessions in which a decrease in depression severity was preceded by the Bond subscale score mostly falling into the Very High range and sometimes the upper end of the High range. There was usually an increase in Goal and/or Task subscale scores and often the Bond subscale score in the same session as the decrease in depression severity (e.g., cases 1, 2, 3, 4, 5, 6, 7, 8, 9). However, sometimes there was a lowering of two or more subscale scores in the same session (e.g., cases 1, 2, 3, 6, 7).

In four case studies there were sessions in which depression severity decreased after a gradual increase in the Bond subscale over two to three sessions (e.g., cases 1, 3, 4, 10). In cases 4, 9 and 10 there was a lowering of depressive symptoms after a building of Goal subscale scores over a couple of sessions. In cases 9 and 10 there was a lowering of depressive symptoms after a building of Task subscale scores over a couple of sessions. In cases 5 and 6 there was a lowering of depression severity after a gradual building of all the subscales over a couple of sessions. Sometimes there was an increase in the subscales scores in the same session, but other times the subscale scores decreased.

In seven cases there were sessions in which an increase in depression severity was preceded by either a gradual building of all subscale scores or all the subscale scores fell in the High or Very High range. There was usually a lowering of Goal and/or Task and sometimes Bond subscale scores in the same session as the increase of depression severity (e.g., cases 1, 2, 3, 6, 7, 8, 10). However, in case 4, there was a session in which an increase in depression severity was preceded by the building of all the subscale scores. The Goal and Task subscale scores continued to increase, but the Bond subscale score decreased in the same session as the rated depression severity. In case 10, the only increase in depression severity was preceded by the building of all subscale scores. Then there was a decrease in all subscale scores in the same session as the rated depression severity.

In seven cases there were sessions in which an increase in depression severity was preceded by one or more subscale scores; particularly the Bond subscale score decreasing (e.g., cases 1, 3, 4, 5, 7, 8, 9). Sometimes there was a decrease in the subscales scores in the same session, but other times the subscale scores increased.

Finally, in four cases there were sessions in which an increase in depression severity was preceded by Goal and/or Task subscale scores decreasing, but the Bond subscale score falling in the Very High range. There was an increase of one or more subscale scores in the same session as the rated depression severity (e.g., cases 5, 6, 7, 9). Usually a decrease in depression severity occurred in the following session.

Overall Summary

There appeared to be five main Goal, Task, and Bond subscales-depression severity patterns that fluctuated throughout the 10 sessions (6 sessions in case 10), but these patterns were not clearly defined and there was no consistency of these patterns within the cases.

Chapter 9 presents a discussion summarizing and integrating the results of the current research, with overall implications and future directions.

Chapter 9: Discussion –Current Research

Overview

The following chapter reviews the aims and research questions of the current research and discusses the implications of the data collected in the context of these. The strengths and limitations of the current research, areas for future research, and the importance of these findings for clients and therapists are also presented.

Review of Aims, Research Questions and Findings.

The overall aim of the current research project was to gain more understanding of the role of the alliance process and symptomatic change within CBT for depression in order to contribute to enhancing the effectiveness of this therapy. The current research also aimed to understand the alliance process and its measurement and its theoretical and empirical underpinnings in order to make careful methodological decisions. Therefore, the current research addressed four key methodological considerations in order to investigate the role of the alliance and symptomatic improvement within CBT for depression. The first methodological consideration of the current research was to review theoretical and empirical support for the alliance within psychotherapy. As detailed in Chapter 2 this review revealed that the role of the alliance is complex and that it varied across theoretical approaches (Raue et al., 1993). It was important to conceptualise the alliance within CBT and align it with the relational quality of the collaborative empiricism (Beck et al., 1979) to determine the most appropriate alliance measure.

The second methodological consideration was to determine the most appropriate alliance scale to measure the alliance within CBT. The current thesis reviewed the four most frequently used alliance scales and then the factor structure of the various versions of the WAI to determine the WAI-SR-O as the most appropriate instrument to measure the alliance and Goal, Task, and Bond components of the alliance within CBT.

The third methodological consideration was to determine a research design that would address the sequential relationship between the alliance and symptomatic improvement and the sequential relationship between Goal, Task and Bond components of the alliance and symptomatic change within CBT for depression. The current research selected a single-

case research design with measurement of alliance and depression severity every session for the first ten sessions of therapy to determine temporal relations.

The fourth methodological consideration was to train raters to rate the WAI-SR-O and establish inter-rater reliability for the current research. In order to increase the inter-rater reliability of the current research, guidelines for rating the Working Alliance Inventory were sought from researchers overseas. These guidelines were modified and expanded to rate the WAI-SR-O within CBT. A rater reliability study was conducted in two stages to provide a forum to train the raters for the current research and establish inter-rater reliability. Stage One investigated whether giving more specific guidelines to rate the WAI-SR-O would lead to less variance in ratings. The results indicated that training rater's well and giving them more opportunity to practice rating sessions were more beneficial. Stage Two investigated whether there was higher inter-rater reliability rating the alliance within CBT if raters had CBT training and/or clinical experience. Findings demonstrated that raters with some CBT training and/or clinical experience may be required to rate the more abstract constructs, such as alliance. The overall inter-rater reliability of the WAI-SR-O in the current research indicated a satisfactory level of agreement. The inter-rater reliability Goal and Bond subscales also indicated a satisfactory level of agreement. The inter-rater reliability of the Task subscale indicated an acceptable, but possible improvable level of agreement. Further discussion of the rating of the WAI SR-O and inter-rater reliabilities is presented later in this chapter

Four research questions were addressed in order to investigate the role of alliance and symptomatic change within CT for depression. First, the current research asked whether a strong early alliance predicted a positive outcome. Second, it questioned whether alliance preceded symptomatic change. Third, the current research explored the intertwined and sequential relationship between alliance and symptomatic change. Finally, the current relationship explored the intertwined and sequential relationship between Goal, Task and Bond components of the alliance and symptomatic change to provide a finer grained analysis of the alliance process and symptomatic change.

Does a strong early alliance predict a positive outcome?

The results of the current research supported previous findings (e.g., Hartley & Strupp. 1983; Henry & Strupp, 1994; Hersoug et al., 2002; Horvath, 2000; Horvath & Luborsky, 1993; Loeb et al., 2006; Kokotovic & Tracey, 1990; Marzali et al., 1999; Reandeau & Wampold, 1991; Strauss et al., 2006) and demonstrated that a strong early alliance is predictive of positive outcome. Total Alliance scores falling into the High or Very High range of the WAI-SR-O could be interpreted as a strong alliance between the therapist and client. In the current research the Total Alliance scores of a majority of the research sample mostly fell in the High range of the WAI-SR-O throughout therapy, with many scores falling into the lower end of the Very High range. There were only five sessions out of 96 rated sessions that the Total Alliance score fell into the upper end of the moderate range. For a majority of the cases the Total Alliance fell in the upper end of the High range or in the lower end of the Very High range in the first session. Furthermore, in a majority of the cases, the Total Alliance score appeared to peak by the third session, which is consistent with the findings of Eaton et al. (1988), Horvath, (1993, 2001), Horvath and Greenberg (1994) and Horvath and Luborsky (1993). The patterns of depression severity of the individual clients across the ten sessions of therapy were mixed. However, the general trend was that most clients experienced a gradual decline in depression severity and with the exception of the client in case 8 they finished therapy with either no depression or mild depression, according to their BDI-II scores. A strong early alliance peaking by the third session and the gradual decline in depression severity in all the cases of the current research does suggest that a strong early alliance is predictive of outcome.

While it is difficult to draw conclusions about the association between the strength of the alliance and premature termination of therapy from one case, the findings of the current research replicated those of other research suggesting that poor early alliance may predict client dropout (Blatt et al., 1996; Bordon, 1979; Botella et al., 2008; Constantino et al., 2002; Kokotovic & Tracey, 1990; Lingiardi et al., 2005; Plotnicov, 1990; Tryon & Kane, 1995). The client in case 10 experienced a gradual decline in depression severity, but terminated her therapy at session 6. For the six sessions the Total Alliance scores fluctuated between the high end of the moderate range to the low end of the High range with the scores ranging between 35.5 and 45 on the WAI-SR-O. The Total Alliance score for the first four sessions mainly fell in the lower end of the High range. The Total Alliance scores appeared to be lower than those of other cases during these early sessions and may not have been 'good enough' for the client and therapist to engage effectively in therapy.

While these lower Total Alliance scores did not appear to hamper her gradual decline in depressive symptoms they may have contributed to her early termination of therapy. In contrast, the dyad in case 8 also experienced a couple of Total Alliance scores falling into the moderate range (i.e., 33.5 and 35.5) during the ten sessions, but this client continued therapy to session 20. However, in this case the Total Alliance scores fell into the upper end of the High range over the first two sessions, dropped to a score of 35.5 in the third session, before developing again over the following sessions. Comparing these two cases demonstrated that if the alliance is strong in the early sessions of therapy it is less likely that the client will prematurely terminate therapy.

It has been reported that client characteristics (Connolly Gibbons et al., 2003; Constantino et al., 2002; Hardy et al., 2001; Hersoung, 2001; Joyce et al., 2003; Joyce & Piper, 1998; Roth & Fonagy, 1996), therapist characteristics (Ackerman & Hilsenroth, 2001, 2003; Barber et al., 2006; Hersoung, 2001; Luborsky et al., 1985 Kivlighan et al., 1998; Mallinckrodt & Nelson, 1991) and environmental factors (Luborsky et al., 1983; Needles & Abramson, 1980) may influence the development and maintenance of the alliance. Any of these factors may have contributed to the dyad in case 10 failing to engage and the client terminating therapy by the 6th session. Notably, the therapist found it difficult to gain rapport with this client and eye contact was minimal. The client was deeply embarrassed about her situation and she became increasingly uncomfortable with the sessions being filmed and talking about her stressful circumstances (see Appendix F for background information). In contrast, clients in the other cases were more easily engaged in therapy. Therefore, a strong early alliance may be necessary to engage a client in therapy to prevent premature termination and various factors (e.g., client, therapist and environmental factors) may influence its development.

Does alliance precede symptomatic change?
In accordance to recent research the current research expected that alliance would precede symptomatic change when symptomatic change is a consequence of specific therapeutic intervention. Pretreatment assessment and initial clinical assessment sessions were not DVD-recorded in the Depression Outcome Study and therefore the interaction between the therapist and client could not be rated in these sessions. However, the depressive severity of seven clients decreased according to their BDI-II scores between the pretreatment assessment and the first session. This may be interpreted that initial alliance building in the way of discussing the clients problems and therapy goals and instilling hope in the client

during the initial assessment session with their therapist may have contributed toward depressive symptoms being alleviated. Hilsenroth and Cromer (2007) maintained that strengthening the alliance during initial assessment was warranted. However, it is possible that other factors, such as the client experiencing a pleasant life experience (Luborsky et al., 1983; Needles & Abramson, 1980) and the client's ability to quickly engage in positive interpersonal relationships (Marmar et al., 1989) may have also contributed to these decreases in depression severity over this time. When Sexton et al. (2005) explored what lay behind an early alliance's prediction of outcome they found that the degree to which the client talked about himself/herself appeared to provide clues about the development of the alliance. They also found that therapist activities, such as active focused listening, keeping the topic centred on the client, providing a relaxed warm atmosphere in-session, allowing the client to become emotionally moved, avoiding the provision of too much information or advice and not using a purely cognitive verbal style promoted important client-therapist connections.

For seven clients (i.e., cases 1, 2, 3, 6, 7, 8, 10) depression severity decreased over the first week (i.e., first two sessions) of therapy. Total Alliance scores over the first week of therapy fell in the higher end of the High range and/or the lower end of the Very High range for a majority of the cases. This may indicate a strong alliance from the start of therapy before any specific therapy interventions have been introduced and suggests that alliance may precede symptomatic change when symptomatic change is a consequence of therapeutic intervention. Alliance-outcome researchers (e.g., Tang & DeRubeis, 1999) did not usually assess the alliance in the first session because this session was deemed to be different from other sessions. However, Busch et al. (2006) found that positive outcomes were predicted by first session gains and concluded that these first sessions of therapy may be relevant for evaluating alliance early in therapy. It has been argued that the cognitive model of CBT would have been demonstrated to the client in the first session and that this could be interpreted as a cognitive intervention (Whisman, 1999). Therefore, cognitive changes may occur in this session and these changes could be considered by some researchers as symptomatic change occurring before the alliance. However, in the current research, clients' rated their decrease in depressive symptoms at the beginning of the first session and so these results could be interpreted as symptomatic change preceding alliance in the first session, but not as a consequence of any specific therapeutic intervention.

In summary, the findings of the current research add to the accumulating evidence on the timing of alliance assessment. The alliance of the pretreatment and initial assessments could not be rated in the current research. Therefore, strong conclusions can not be drawn as to whether alliance preceded symptomatic change prior to the first therapy session. Potentially, the alliance may have contributed to the decrease in depression severity experienced by a majority of the clients between the first pretreatment assessment and the first therapy session. Furthermore, for nine cases the alliance was strong in the first session which could have suggested alliance building prior to that session. Client rated symptomatic change at the beginning of the first session suggested that their depressive symptoms were alleviated before any specific therapeutic intervention. Together, the results of the current research could suggest that a strong early alliance may precede symptomatic change when symptomatic change is a consequence of specific therapeutic intervention. However, causal direction (if any) has not been clearly established. Barber et al., (2000) suggested that it is important to not only query whether the alliance during early therapy predicts outcome or whether alliance predicts early symptomatic change, but to address what is the intertwined and sequential relationship between alliance and client improvement. From the results of the current research it could be speculated that reciprocal relations between the alliance process and symptomatic change may start in the assessment stage of therapy (e.g., cases 2, 3, 4, 5, 8, 9, 10).

What is the intertwined and sequential relationship between alliance and symptomatic change?

For all the cases, results of the current research suggested a clear reciprocal effect between Total Alliance scores and depression severity in some sessions during therapy, but in other sessions the reciprocal effect was not as clear. There appeared to be four different alliance-depression severity patterns that fluctuated in a random manner throughout the sessions and the random manner in which the alliance-depression severity patterns occurred was different for each case. In all the cases the Total Alliance score mostly fell into the upper end of the High range or the lower end of the Very High range in the preceding session and in the same session as the rated depression severity in all four patterns. While these Total Alliance scores followed a linear pattern within this High/Very High range band, they were not static and they fluctuated either each session or weekly. Therefore the alliance appeared to be in a constant state of flux ranging from development, rupture or repair. Furthermore, as demonstrated by examples of in-session content of the cases in the results section of the current research (Chapter 8), various client, therapist and environmental

factors, such as a life crisis or barriers affecting the completion of a homework task, may have had an influence on which alliance-depression severity pattern occurred in which session. This suggests that the individual circumstances of the client and therapist-client interaction may have dictated the various alliance-depression severity patterns and when they occurred during therapy.

In all ten cases there were sessions in which there was a decrease in depression severity and an increase in the Total Alliance score in the same session. This could be interpreted as a strong alliance from the previous session contributing to a lowering of depressive symptoms and subsequently an increase in alliance in the same session. A clear reciprocal relationship was demonstrated between the Total Alliance scores and depression severity over the two sessions. During this alliance-depression severity pattern it is possible that the client may not have experienced any life crises out of session or if they did experience a life crisis they had skills to be able to cope with the stressor. They may have been able to complete their set homework task and in doing so develop more skills and/or enjoyment in life. Furthermore, the therapist and client may have effectively worked on the client's goals and problems and/or collaboratively designed and assigned the client's next homework task. To fully understand what has influenced the occurrence of the different alliance-depression severity patterns, the current research argues that in-session content, environmental factors and other therapeutic processes, such as homework would need to be investigated.

In nine cases there were sessions in which there was an increase in depression severity and a decrease in the Total Alliance score in the same session. When this pattern occurred there was usually a gradual building of the Total Alliance scores over two to three sessions preceding the increase in depression severity. Mixed alliance–depression severity patterns occurred in the following session, but in a majority of the cases there was a decrease in depression severity and an increase in Total Alliance scores. The gradually developing alliance was strong at the time of the increase in depression severity and it is possible that this high quality of alliance may have contributed to the increase in depression severity scores only lasting for the one session. This alliance-depression severity pattern may have been influenced by environmental, in-session content and/or other therapeutic processes. Decreases in the alliance can reveal mistimed interventions, the clients reaction to the shifting techniques of the therapist, the therapist failing to understand the clients emerging resistance toward interventions, the client adapting to trialing and completing activities

outside of therapy, the therapist not attending to or responding ineffectively to ruptures, and /or the client may feel uncomfortable or anxious as he/she starts to look at new ways of behaving and thinking (A. T. Beck et al, 1979; J. S. Beck, 1995). It is also possible that the client may have experienced a life crisis outside of the therapy session and/or experienced barriers completing homework which may have increased his/her level of depression severity at the beginning of the session and subsequently the alliance to be lowered in the session. Alternatively the client may have brought up a particularly disturbing or painful issue (e.g. past loss) to work on in-session and the client and therapist may struggle with this issue until it is resolved for the client or at least it was resolved enough for the client to feel less distressed. Furthermore, the therapist and client may have effectively worked on the crisis and or a rupture in the alliance during the session, but a decrease the depression severity and subsequently an increase in alliance may not have been evident until the following session. The reciprocal relationship between the Total Alliance scores and depression severity was not as clear as that of the first pattern and it usually occurred over two to three sessions. However, it could be hypothesized that an increase in depression severity lowers the alliance, but if the quality of the alliance is high enough (i.e., from the gradual building of Total Alliance scores in previous sessions) for the client and therapist to work effectively on the crisis or rupture in-session (Horvath & Greenberg, 1994), there may be a decrease in depression severity and then an increase in the alliance in the following session.

In all ten of the cases there were sessions in which there was a decrease in depression severity and a decrease in the Total Alliance score in the same session. Sometimes, in the following session the depression severity continued to decrease and the Total Alliance score usually increased. It could be hypothesized that a strong alliance in the previous session precipitated a lowering in depression severity and this quality of alliance and/or decrease in depression severity may have contributed to the increase in Total Alliance score in the following session. This alliance-depression severity pattern may have been influenced by environmental factors, in-session content and/or other therapeutic processes as already discussed in the first two alliance-depression severity patterns. Again the reciprocal effect between the Total Alliance scores and depression severity was not clear and it occurred over two to three sessions.

In seven of the cases there were sessions in which there was an increase in depression severity and an increase in the Total Alliance score in the same session. Mixed alliance–

depression severity patterns occurred in the following session including: decrease in both depression severity and Total Alliance scores, increase in both the depression severity and Total Alliance scores, decrease in depression severity and an increase in Total Alliance scores and an increase in depression severity and a decrease in Total Alliance scores. This alliance-depression severity pattern may also have been influenced by environmental factors, in-session content and other therapeutic processes as already discussed in the first two alliance-depression severity patterns. The reciprocal effect between the Total Alliance scores and depression severity was not clear because there was no main alliance-depression severity pattern occurring in the following session.

There is only one known study (i.e., Barber et al., 2000) that has indirectly investigated the reciprocal effects of the alliance and outcome. The methodology of that study did not allow the researchers to measure these variables every session, but they concluded that alliance and outcome may be so intricately intertwined that they may amplify each other rapidly. The current research demonstrated four different alliance-depression severity patterns which fluctuated every session in a random manner throughout the therapy sessions. There appeared to be a clear reciprocal relationship between alliance and depression severity in some sessions, in that, a strong alliance may have contributed to a decrease in depression severity which subsequently increased the alliance. However, other reciprocal patterns were not as clear and they may have taken place over two to three sessions. For example, an increase in depression severity may have lowered the alliance, but if the alliance had been gradually building in the previous sessions, depression severity could decrease and alliance could increase in the following session. To fully understand these quickly occurring shifts in the alliance and symptomatic change and what influences the occurrence of the different alliance-depression severity patterns, environmental factors (e.g., life crisis), in-session content and other therapeutic processes (e.g. homework processes) could be investigated. The different alliance–depression severity patterns occurring throughout the cases may reflect the individual and varied circumstances and beliefs of each client and varied client-therapist interactions. It was expected by the current research that closer inspection of the alliance at the subscale level would reveal deeper understanding of the reciprocal relationship between the alliance and symptomatic change.

What is the intertwined and sequential relationship between Goal, Task and Bond components of the alliance and symptomatic change?

Most researchers assumed that the alliance is a one-factor construct (Hatcher & Barends, 1996). However, Andrusyna et al. (2001) argued that the alliance was far more complex than many first believed and suggested that the components of alliance needed to be explored as independent constructs. Research findings on this matter appeared to be inconsistent (Horvath & Greenberg, 1986; Horvath & Marx, 1990).

In most cases of the current research the Bond subscale was the most developed (i.e., highest subscale scores) throughout the ten sessions. Furthermore, for the early sessions of therapy the scores of the Bond subscales fell in the High or Very High range. However, in the first session of case 10 the Bond subscale score fell into the moderate range (i.e., score of 12 on the Bond subscale). It increased over sessions 2-5, but fell into the lower end of the moderate range in the 6th session (i.e., score of 10.5 on the Bond subscale). Principe et al., (2006) found that a client's return for a second session was significantly related to the Bond subscale of the WAI-S-C. The client in case 10 was diagnosed with generalized anxiety disorder comorbid (on the CIDI) with her depressive symptoms. The quality of the Bond may not have been high enough to quell her anxiety and permit her to engage in tasks or pursue goals (Pimsoff, 1994). In session 3 the client and the therapist failed to collaboratively agree on goals and what was important for the client to work on. From these results it is suggested that the quality of the Bond component of the alliance early in therapy needs to be high enough to allow the client and the therapist to work together effectively and to prevent premature termination.

The current research explored the reciprocal effects between the alliance and symptomatic change on a deeper subscale level and expected to gain a more detailed analysis about the role of alliance process and of symptomatic change within CBT. However, while there were some main trends, the Goal, Task and Bond subscale-depression severity patterns were not as clearly defined on this level. Therefore, direction of the temporal relations between the Goal, Task and Bond subscales and depression severity were not as apparent as reciprocal effects between the Total Alliance score and depression severity. There appeared to be five main temporal patterns of Goal, Task, and Bond subscales and depression severity that fluctuated throughout the sessions and the random manner in which these patterns occurred was different for each case. For a majority of the cases there were sessions in which a decrease in depression severity was most often preceded by the

Bond subscale score mostly falling into the Very High range and sometimes the upper end of the High range. Subsequently, there was usually an increase in Goal and/or Task subscale scores and often the Bond subscale score in the same session as the decrease in depression severity. Therefore, a strong Bond, where the client feels appreciated and not judged and there is mutual liking and respect between the client and the therapist, may facilitate a decrease in depression severity. This lowering of depressive symptoms may then contribute to the client and therapist collaboratively agreeing on goals and working together on tasks. It may also promote further warmth, trust and friendliness between the client and therapist.

For the majority of the cases there were sessions in which an increase in depression severity was preceded by either a gradual building of all subscale scores or all the subscale scores falling in the High or Very High range. There was generally a lowering of the Goal and/or Task and sometimes Bond subscale scores in the same session as the increase of depression severity. This increase in depression severity at the beginning of the session may be due to environmental factors outside of the therapy session and/or the client experiencing barriers completing homework, and may subsequently influence clarity of goals and tasks for that session. Furthermore, there were sessions in which an increase in depression severity was preceded by one or more subscale scores, particularly the Bond subscale score decreasing. In this preceding session the client may have felt judged and perceived that he/she was not appreciated as a person by the therapist and the client and therapist may have struggled to get clarity about goals and/or tasks and/or what is right for the client. It was difficult to determine the bidirectional relationship in this particular pattern because sometimes there was a decrease in subscale scores and other times there was an increase in subscale scores in the same session as the increase of depression severity.

Definite conclusions could not be drawn about the reciprocal effects between the Goal, Task and Bond subscales and depression severity. However, the findings of the current research appear to be consistent with the results of Rector et al. (1999), in that symptom change is greater in the context of a strong bond between the client and therapist. Rector et al. (1999) could not determine bidirectional relationships between the alliance components and cognitive change in their study and therefore they could only demonstrate that agreement between the client and therapist on the goals and tasks of therapy predicted change in dysfunctional beliefs. Similar to Rector et al. (1999), the current research found

that in some sessions of some cases, agreement or lack of agreement between the client and therapist on the goals and tasks in the previous session may have contributed to a decrease or increase in depressions severity respectively. Goal, Task and Bond subscale-depression severity patterns in the current research were not clearly defined, but there was some evidence of Goal, Task and Bond subscale scores increasing following a decrease in depression severity and decreasing following an increase in depression severity in the same session that the depression severity was rated. In summary, in some sessions in some cases the current research demonstrated a unidirectional relationship between the Goal, Task and Bond components of the alliance and depression severity and in some sessions in some cases there was evidence of a bidirectional relationship. Similar to the exploration of reciprocal effects between the alliance and depression severity the lack of clarity surrounding reciprocal effects between Goal, Task and Bond components of the alliance and depression severity may be due to individual personality, circumstances, comorbidity and belief systems of the client and unique therapist-client interactions.

Strengths of the Current Research

This research has a number of strengths. The alliance construct was conceptualized within CBT to determine an appropriate alliance measure. Raters were thoroughly trained to rate the WAI-SR-O and inter-rater analyses were conducted. While there was room for improvement, the overall inter-rater reliability of the WAI-SR-O and the Goal and Bond subscales were at a satisfactory level. A single-case research design was utilized with multiple ratings of alliance and depression severity which allowed the sequential relationship between these two variables to be demonstrated on an individual level. The current research utilized data from The Depression Outcome Study conducted at The School of Psychology, Massey University, Albany (2006-2009). Novice therapists are usually considered a limitation. They are generally trained to a standard, can fail to use the CBT model and techniques in a flexible way, tend to follow manualised protocols, such as the homework protocol more rigidly than experienced therapists and do not have as much expertise to repair ruptures in the alliance. The three therapists of The Depression Outcome Study were trained to deliver Beckian CBT (Beck et al., 1979) and utilize the guiding model for reviewing, designing and assigning homework (Kazantzis, MacEwan, et al., 2005). However, despite juggling the delivery of CBT interventions and reviewing, designing and assigning homework according to the homework protocol, and attending to the alliance between themselves and the client, the Total Alliance scores were mostly in the High range or lower end of the Very High range for the current research. They were

closely supervised by an experienced CBT clinician. Furthermore, the CTS (Young & Beck, 1980) was utilized to provide standards for quality control monitoring and therapist and client safety and treatment integrity. The HAACS (Kazantzis, Wedge et al., 2005) was utilised to determine the extent of therapist adherence and competence in the design, assignment, and review of homework as described in the homework protocol.

Limitations of the Current Research and Future Recommendations

The current research has a number of limitations. However, this was a good first attempt to investigate the reciprocal effects of alliance processes and symptomatic change within the constraints of time, and advances what is currently known about in this area. The sample size was small and findings may not generalize to other populations and/or theoretical orientations. The guidelines were specifically developed to facilitate the rating of the WAI-SR-O within CBT, as was the training of the raters. The briefer WAI-SR-O was utilized to assess the alliance rather than the longer original form of the WAI. It is possible that the longer original 36 item WAI may have more adequately captured the collaborative process between the client and the therapist. However, raters were rating three other scales alongside the 12 item WAI-SR-O. Time limitations due to limited finances, the amount of information to be processed and the extra items to rate may have made it more difficult to rate every item of the longer scale effectively and there would have been more room for discrepancy. The alliance was assessed only from the observer's perspective and no therapist and client data were collected which would have allowed for comparison of alliance ratings. There were only two raters. The use of multiple raters (e.g., three raters) may have increased rater reliability and given added protection against raters drift. Furthermore, raters were not specifically chosen for the current research. The rater reliability study conducted within the current study demonstrated the need to select raters with some clinical experience. Only one of the raters in the current research had some clinical experience. In accordance with previous research methodology the current research only utilized one outcome measure, the BDI-II. However, the addition of other outcome scales, such as the ASQ and the SOFAS may have given a more balanced investigation of the role of symptomatic change within CBT. The use of recalibration training sessions after the completion of 10 consecutive ratings also offers protection against raters drift. Recalibrations for the current research ceased after the first three sets of 10 sessions. Continued recalibrations and feedback to the raters may have yielded high inter-rater reliability. As already demonstrated in the rater reliability study rating the Task subscale was problematic. However, inter rater reliability improved in the third set of recalibrations

after feedback from the first two sets of recalibrations. Without continuous feedback the overall inter-rater reliability for the Task subscale decreased, whereas the overall inter-rater reliability of the Goal and Bond subscales increased. Inter-rater reliability was lower that that obtained by Hatcher and Gillaspy (2006). However, Hatcher and Gillaspy (2006) had larger samples and therapists were of mixed theoretical orientations. Notably there was less variability in the scoring range of the Task subscale (see figure 8.4). The therapists of the Depression Outcome Study adhered very closely to the newly developed homework protocol (Kazantzis, MacEwan, et al., 2005) to design and assign homework and this may have influenced inter-rater reliability of the Task subscale. This adherence was monitored by the Depression Outcome Study coordinators and any non adherence was dicussed in supervision. While an atmosphere of collaboration was fostered between the therapist and the client to design and assign homework projects, the protocol involved the therapist and client agreeing on how the homework task could be completed, when it was possible for the client to complete the task and how often it was necessary for the client to complete the task. It is possible that the therapist's rigid adherence to the protocol may have caused a lack of variability in the scoring range of the Task subscale. The administration and completion of homework usually involves collaborative processes (A. T. Beck et al., 1979; J. S. Beck, 2005; Blackburn & Davidson, 1995; Coon et al., 2005; Freeman & Rosenfield, 2002; Tompkins, 2003). It is also possible that there is some overlap between alliance and homework processes within CBT which would need further investigation. The three therapists were selected, rigorously trained and closely supervised and so the results from the current research may not generalize to more naturalistic clinical settings, where the range of therapist competence is likely to vary widely and where there might be little or no emphasis on therapist adherence to protocol. A final limitation is that third-variable confounds, such as client and therapist factors, comorbidity, and environmental, social, economic and legal factors, can not be ruled out within a single-case research design.

Clear conclusions about the reciprocal effects between the alliance and depression severity and the reciprocal effects between the Goal, Task and Bond components of the alliance and depression severity could not to be drawn from the findings of the current research. Individual variance between and within each case may contribute to these results. It is recommended that further research be conducted to extend the findings of this study.

In the first instance it is recommended that the current research be replicated utilizing a larger sample and statistical techniques such as growth curve modeling, growth mixture

modeling and dynamical systems modeling which are capable of determining the mechanisms underlying change (Laurenceau et al., 2007). Reciprocal effects between the alliance process and symptomatic change may start as early as the initial assessment and therefore, it may be beneficial to extend alliance measurement to cover pretreatment assessments. In light of the findings of the rater reliability study conducted within the current research it is recommended that future researchers choose their raters carefully and employs those with clinical knowledge and experience. This may enhance inter-rater reliability (Burstein & Carkhuff, 1968; Moras & Hill, 1991).

It is proposed that it would be beneficial to investigate the role of the alliance process from the perspective of the client. While observers' data usually permits replication and yields information that is more objective, observer versions of alliance instruments are not usually recommended because observers must make inferences about clients' thoughts, feelings and motives for which evidence can be more difficult to find. What appears to the outside observer to be a poor alliance may be experienced by the client as a good alliance and visa versa. While client rated alliance can be biased (e.g., halo effect), research has shown client and observer evaluations to be stronger than therapist judgment (Horvath & Symonds, 1991).

The current research only utilized one outcome measure: the BDI-II. It would also be beneficial to evaluate change in dysfunctional beliefs and functioning with the ASQ and SOFAS to get an overall picture of symptomatic change.

It has been reported that the alliance between the client and therapist are different in early, middle and late therapy. Future research may like to extend the current research and to gain further understanding of the role of the alliance process and symptomatic change through middle and late therapy.

Future research could utilize the original 36 item WAI which may capture the collaboration process between the client and therapist more adequately than shorter versions of the WAI. The current research was limited to only four items per subscale. However, as discussed earlier in the discussion session of Chapter 7 it may be more appropriate for future researchers to investigate the use of alternative alliance scales or develop a measure that will adequately capture the collaborative empiricism of CBT.

Total Alliance scores and Goal, Task and Bond subscale scores of the WAI can only look at shifts between sessions, rather than within them and it is possible that that this alliance scale is missing a critical part of the change process within CBT. Eyeballing the raw data of rated subscale items allows some understanding of what is happening inside each session. For example, respect and mutual liking between the client and therapist may be high, but they may struggle to agree on goal and tasks of the therapy. Future researchers may like to extend the current research to investigate what is happening between sessions (e.g., environmental factors), in-session, 'third variables' (e.g. client and therapist factors) and other therapeutic processes, such as homework processes to explore this area further and provide more clarity on the reciprocal effects between alliance and symptomatic change. Separately alliance and homework processes have received much attention within CBT. However, few studies (e.g., Carroll et al., 2005; DeRubeis & Feeley, 1990; Dunn et al., 2006; Feeley et al., 1999; Mueser et al., 2008; Taft et al., 2003; Woody & Adessky, 2002) have investigated the relationship between alliance and homework and the results are unclear. Homework is assigned from the first session and this may present an immediate challenge for the alliance (Dunn, Morrison & Bentall, 2002, 2006; Fehm & Kazantzis, 2004; Kazantzis, Lampropoulos, & Deane, 2005). Therefore, it is important to understand how homework can be assigned without straining the alliance. A qualitative descriptive study may be more appropriate to deepen the understanding of these underlying "third variable" factors (e.g., Colli & Lingiardi, 2009; Elvins & Green, 2008).

Furthermore, it remains to be seen to what extent the current results are applicable to other forms of psychotherapy for a variety of disorders and not for depression within CBT alone.

Implications for Clinical Practice

There are a number of clinical implications of the current research. The results underscore the importance of therapists attending to the alliance process within therapy and understanding the complex sequential relationship between alliance and symptomatic change. The findings of this research may have implications for the practice of CBT and for training and supervising therapists to work effectively with the individual client. It may be important in early therapy to focus more on strategies to engage clients in therapy and develop the alliance rather than focusing on technical competence. Wampold (2001) maintained that the alliance accounted for more variability than the totality of specific therapeutic interventions. Therapists working with clients who exhibit lower ability to engage in a positive interpersonal relationship may need to seek alternative ways of

developing an adequate alliance in the early stages of therapy; failure to do so could precipitate premature termination. In the current research there appeared to be four different alliance-depression severity patterns that fluctuated in a random manner throughout the sessions and random manner in which the alliance-depression severity patterns occurred was different for each case. It was demonstrated that it was important to develop a strong alliance and maintain it throughout therapy to lessen depression severity and this lowering of depressive symptoms may increase the alliance between the client and therapist. However, the CBT therapist needs to be mindful that each client-therapist interaction is unique and to be sensitively attuned to each client's personality, circumstances and belief systems (A. T. Beck et al., 1979). Furthermore, while conclusions could not be drawn to determine reciprocal effects between the Goal, Task and Bond components of the alliance and depression severity there was evidence to suggest that CBT students may need to acquire clinical skills to strengthen the Bond component of the alliance including: gaining respect and liking from the client, exhibiting unconditional positive regard and appreciating the client as a person, to lower depressive symptoms and to facilitate collaboration and agreement about the goals and tasks of therapy (Goldfried, 2007, Norcross, 2002, Pinsoff, 1994). Principe et al. (2006) demonstrated that a client's return for a second session was significantly related to the Bond subscale of the WAI-S-C. Together, these results inform treatment to enhance recovery of depressive symptoms.

Conclusions

In conclusion, the current research confirmed and extended previous reports about temporal relations between the alliance process and symptomatic change within CBT. The single-case research design with multiple assessments offered rich information about individual variability and closer detailed analysis which has not been demonstrated in previous studies. The current research had several methodological limitations (e.g., the employment of only two raters of which only one had some clinical experience, the utilization of an observer alliance scale which made rating such a complex and subjective construct as the alliance more difficult, unclear reasons as to why the inter-rater reliability of the Task subscale was problematic, concern as to whether the the briefer form of the WAI (i.e. WAI-SR-O) or even the original 36 item WAI adequately captures the collaboration process between the client and the therapist within CBT and the utilization of a single-case research design which made the ruling out of underlying confounding factors and the generalization of the results to other populations and psychotherapy orientations difficult) and therefore the results must be interpreted cautiously.

Results demonstrated that a strong early alliance is a predictor of a positive outcome and that poor early alliance may lead to premature termination of therapy. The current research demonstrated that depression severity decreased between the pretreatment assessment and the first therapy session and then again in the first week of therapy for a majority of the clients. However, while the findings of the current research add to the accumulating evidence on the timing of alliance assessment it was difficult to draw definite conclusions as to whether alliance precedes symptomatic change prior to the first therapy session. Client rated symptomatic change at the beginning of the first session suggested that their depressive symptoms were alleviated before any specific therapeutic intervention. A strong alliance in the first session could suggest that initial alliance building in the way of discussing therapy goals and instilling hope in the client may have taken place during the initial assessment session. Therefore, the results of the current research could indicate that a strong early alliance precedes symptomatic change before specific therapeutic interventions are introduced. Furthermore, it is possible that reciprocal relations between the alliance process and symptomatic change may start in the assessment stage of therapy.

The current research also addressed the intertwined and sequential relationship between alliance and client improvement. Novel findings have demonstrated some reciprocal effects between the alliance and symptomatic change and highlight the important role that both the alliance and client improvement play within therapy. The current research demonstrated four different alliance-depression severity patterns which fluctuated every session in a random manner throughout the therapy sessions. It was difficult to draw definite conclusions about the reciprocal effects between the alliance and depression severity. However, there was evidence of a clear reciprocal relationship in some sessions in some cases. A strong alliance in the previous session may have contributed to a decrease in depression severity at the beginning of the session which subsequently increased the alliance for that session. However, in other sessions reciprocal patterns were not as clear. To fully understand what influences the occurrence of the different alliance-depression severity patterns during therapy, individual environmental factors, in-session content, client and therapist factors and other therapeutic processes could be investigated. The different alliance–depression severity patterns occurring throughout the cases may reflect the individual and varied circumstances and belief systems of each client and unique client-therapist interactions.

Given the complexity of the alliance process the current research also addressed the reciprocal effects of the alliance subscales and symptomatic change. Definite conclusions could not be drawn about the reciprocal effects between the Goal, Task and Bond subscales and depression severity because Goal, Task, and Bond subscale-depression severity patterns were ambiguous. However, there was some evidence that symptomatic change was greater in the context of a strong bond between the client and therapist. Likewise, there was some evidence of Goal, Task Bond subscale scores increasing following a decrease in depression severity and decreasing following an increase in depression severity in the same session that the depression severity was rated. In some sessions in some cases the current research demonstrated a unidirectional relationship between the Goal, Task and Bond components of the alliance and depression severity and in other sessions there was evidence of a bidirectional relationship. Similar to the exploration of reciprocal effects between the alliance and depression severity the lack of clarity surrounding reciprocal effects between Goal, Task and Bond components of the alliance and depression severity may reflect individual personality, circumstances, comorbidity and belief systems of the client and unique therapist-client interactions.

In summary, utilizing data collected in every session has allowed the current research to derive cautious inferences about the reciprocal relationship between the alliance process and symptomatic change. However, the random manner in which Total Alliance-depression severity patterns and Goal, Task and Bond subscale-depression severity patterns occurred throughout the rated therapy sessions and the variance of these patterns within each case has suggested that these reciprocal effects may be determined by individual client factors, therapist factors, unique client-therapist dyads and environmental factors. These results further illustrated the complexity of the alliance process within psychotherapy. The current research has offered preliminary findings about the interplay between the alliance process and symptomatic change. It provided strong evidence that a strong early alliance predicted a positive outcome and was needed to engage the client and prevent premature termination of therapy. Further investigation is warranted in this important area of process research.

References

Ackerman, S. J., & Hilsenroth, M. J. (2001). A review of therapist characteristics and techniques negatively impacting the therapeutic relationship. *Psychotherapy: Theory, Research, Practice, Training, 38,* 171-185.

Ackerman, S. J., & Hilsenroth , M. J. (2003). A review of therapist characteristics and Techniques positively impacting the therapeutic relationship. *Clinical Psychology Review, 23,* 1-33.

Ackerman, S. J., Hilsenroth, M. J., Baity, M. R., & Blagys, M. D. (2000). Interaction of therapeutic process and alliance during psychological assessment. *Journal of Personality Assessment, 75,* 82-109.

Adler, L. A., Spencer, T., Faraone, S. V., Reimherr, F.W., Kelsey, D., Michelson, D., & Biederman, J. (2005). Training raters to assess Adult ADHD: Reliability of ratings. *Journal of Attention Disorders, 8,* 121-126.

Agras, W. S., Rossiter, E. M., Arnow, B., Schneider, J. A., Telch, C. F., Raeburn, S. D., Bruse, B., Perl, M., & Koran, L. M. (1992). Pharmacological and cognitive- behavioral treatment for bulimia nervosa: A controlled comparison. *American Journal of Psychiatry, 149,* 82-87.

American Psychiatric Association (2000). *Diagnostic and statistical manual of mental health disorders* (4th ed, text revision). Washington D.C.: American Psychiatric Association.

Andrews, G., & Peters, L. (1998). The psychometric properties of the Composite International Diagnostic Interview. *Social Psychiatry and Psychiatric Epidemiology, 33,* 80-88.

Andrusyna, T. P., Tang, T. Z., DeRubeis, R. J., & Luborsky, L. (2001). The factor structure of The Working Alliance Inventory in cognitive-behavioral therapy. *Journal of Psychotherapy Practice and Research, 10,* 173-178.

Arnhoff, F. N. (1954). Some factors influencing the unreliability of clinical judgments. *Journal of Clinical Psychology, 10,* 272-275.

Athey, T. R., & McIntyre, R. M. (1987). Effect of rater training on rater accuracy: Levels of processing theory and social facilitation theory perspectives. *Journal of Applied Psychology, 72,* 567-572.

Bachelor, A. (1991) Comparison and relationship to outcome of diverse dimensions of the helping alliance as seen by client and therapist. *Psychotherapy, 28,* 534-549.

Bandura, A. (1969). *Principles of behavior modification.* New York: Holt, Rinehart and Winston, Inc.

Bannister, B. D., Kinicki, A. J., Denisi, A.S., & Horn, P.W. (1987). A new method for the statistical control of rating error in performance ratings. *Educational and Psychological Measurement, 47,* 583-596.

Barber, J. P., Connolly, M. B., Crits-Christoph, P., Gladis, L. & Siqueland, L. (2000). Alliance predicts patient's outcome beyond in-treatment change in symptoms. *Journal of Consulting and Clinical Psychology, 68,* 1027-1032.

Barber, J.P., Gallop, R., Crits-Christoph, P., Barrett, M. S. Klostermann, S., McCarthy, K. S., & Sharpless, B.A. (2008). The role of alliance and techniques in predicting outcome of supportive-expressive dynamic therapy for cocaine dependency. *Psychoanalytic Psychology, 25,* 461-482.

Barber, J. P., Gallop, R., Crits-Christoph, P., Frank, A., Thase, M. E.., Weiss, R. D., & Connolly Gibbons M., (2006) The role of therapist adherence, therapist competence, and alliance in predicting outcome of individual drug counseling: Results from the National Institute Drug Abuse Collaborative Cocaine Treatment Study. *Psychotherapy Research, 16,* 229-240.

Barber, J. P., Liese, B. S., Abrams, & M. J. (2003). Development of the cognitive therapy adherence and competence scale. *Psychotherapy Research, 13,* 205-221.

Barber, J. P., Luborsky, L., Crits-Christoph, P., Thase, M. E.., Weiss, R. D., Frank, A., Onken, L., & Gallop, R. (1999). Therapeutic alliance as a predictor of outcome in treatment of cocaine dependence. *Psychotherapy Research, 9,* 54-73.

Barber, J. P., Luborsky, L., Gallop, R., Crits-Christoph, P., Frank, A., Weiss, R., Thase, M. E., Connolly, M., Gladis, M., Foltz, C., Siqueland, L. (2001). Therapeutic alliance as predictor of outcome and retention in the NIDA Collaborative Cocaine Treatment Study. *Journal of Consulting and Clinical Psychology, 69,* 119-124.

Barlow, D.H., Gorman, J. M., Shear, M.K., & Woods, S.W. (2000). Cognitive-Behavioral therapy, imipramine, or their combination for panic disorder: A randomized controlled trial. *JAMA, 283,* 2529-2536.

Barlow, D. H., & Hersen, M. (1984). *Single case study experimental designs: Strategies for studying behavior change* (2nd ed.). Boston: Allyn & Bacon.

Bartko J. J., & Carpenter, W. T. (1976). On the methods and theory of reliability. *Journal of Nervous and Mental Disease, 163,* 307-317.

Beck , A. T. (1976). *Cognitive therapy and the emotional disorders.* New York: International University Press.

Beck, A. T., Freeman, A., Davis, D. D. & Associates (2004). *Cognitive therapy of personality disorders* (2nd ed.). New York: The Guilford Press.

Beck, A. T., Rush, A. J., Shaw, B. E., & Emery, G. (1979). *Cognitive therapy of depression.* New York: The Guilford Press.

Beck, A. T., Steer, R. A., & Brown, G. K. (1996). *Manual for the BDI-II.* San Antonia, TX: The Psychology Corporation.

Beck , A. T., & Weishaar, M. E. (1989). Cognitive therapy. In R. Corsini & D. Wedding (Eds.), *Current psychotherapies* (4th ed. pp. 285-320) Ithasca, Illinois: F. E. Peacock Publishers, Inc.

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

Been, H., & Winston, A. (1998). The therapeutic alliance in brief psychoanalytic psychotherapy: A drive-conflict perspective. In J. D. Safran & J. C. Muran (Eds.), *The therapeutic alliance in brief psychotherapy* (pp. 15-37). Washington, DC: American Psychological Association.

Begental, J. F.T., & McBeath, B. (1995). Depth existential therapy: Evolution since World War II. In B. Bongar & L. E. Beutler (Eds.), *Comprehensive textbook of Psychotherapy: Theory and practice* (pp. 111-122). New York: Oxford University Press.

Begental, J. F.T., & Sterling, M. M. (1995). Existential-humanistic psychotherapy: New perspectives. In A. S. Gurman & S. B. Messer (Eds.), *Essential psychotherapies: theories and practice* (pp. 226-260). New York: Guilford Press.

Benjamin, & Karpak (2001). Personality disorders. *Psychotherapy, 38,* 487-491.

Bernardin, H. J. (1977). Behavioral expectation scales vs. summated scales: A fairer comparison. *Journal of Applied Psychology, 62,* 422-428.

Bernardin, H. J. (1978). Effects of rater training on leniency and halo errors in student ratings of instructors. *Journal of Applied Psychology, 63,* 301-308.

Bernardin, H. J., Alvares, K. M., Cranny, C. J. (1976). A recomparison of behavioral expectation scales to summated scales. *Journal of Applied Psychology, 61,* 564-570.

Bernardin, H. J., & Pence, E. C. (1980). Effects of rater training: Creating new Response sets and decreasing accuracy. *Journal of Applied Psychology, 65,* 60-66.

Bernardin, H. J., & Walter, C. S. (1977). Effects of rater training and diary keeping on psychometric error in ratings. *Journal of Applied Psychology, 62,* 64-69.

Beutler, L., Malik, M., Alimohamed, S., Harwood, T., Talebi, H., Noble. S., & Wong, E. (2004). Therapist variables. In M. J. Lambert (Ed.), *Bergin and Garfield's Handbook of psychotherapies and behavior change* (5th ed., pp 227-306). New York: John Wiley & Sons.

Blackburn, I. M., & Davidson, K. M. (1995). *Cognitive therapy for depression and anxiety.* Oxford: Blackwell Science Ltd.

Blackburn, I. M., James, I. A., Milne, D. L., Baker, C., Standart, S. Garland, A., & Reichelt, F. K. (2001). The Revised Cognitive Therapy Scale (CTS-R): Psychometric properties. *Behavioural and Cognitive Psychotherapy, 29*, 431-446.

Blackburn, I. M., & Twaddle, V. (1996). *Cognitive therapy in action: A practitioners casebook.* London: Souvenir Press.

Blampied, N. M. (1999). A legacy neglected: Restating the case for single-case research in cognitive behavior therapy. *Behavior Change, 16*, 89-104.

Blampied, N. M., (2001). The third way: Single-case research, training, and practice in Psychology. *Australian Psychology, 36,* 157-163.

Blatt, S. J., Zuroff, D. C., Quinlan, D. M., & Pilkonis, P. A. (1996). Interpersonal factors in brief treatment of depression: Further analyses of the National Institute of Mental Health Treatment of Depression Collaborative Research Program. *Journal of Consulting and Clinical Psychology, 64,* 162-171.

Bond, M., Banon, E., & Grenier, M. (1998). Differential effects of interventions on the therapeutic alliance with patients with personality disorders. *The Journal of Psychotherapy Practice and Research,7, 301-318.*

Bordin, E. S. (1979). The generalizability of the psychoanalytic concept of the working alliance. *Psychotherapy: Theory, Research and Practice, 16,* 252-260.

Bordin, E. S. (1994) Theory and research on the therapeutic working alliance: New Directions. In A. O. Horvath & L. S. Greenberg (Eds.), *The working alliance: Theory, research and practice* (pp. 13-37). New York: John Wiley & Sons.

Borman, W. C. (1975). Effects of instructions to avoid halo error on reliability and validity and performance evaluation ratings. *Journal of Applied Psychology, 60,* 556-560.

Borman, W. C. (1979). Format and training effects on rating accuracy and rater errors. *Journal of Applied Psychology, 64,* 410-421.

Bottella, L. Corbella, S., Belles, L., Pacheco, M., Gomez, A. M., Herrero, O., Ribas, E., & Pedro, N. (2008). Predictors of therapeutic outcome and processes. *Psychotherapy Research, 18,* 535-542.

Bowlby, J. (1982). *Attachment* (2nd ed.).United States of America: Basic Books, A Division of Harper Collins Publishers.

Bowlby. J. (1988). *A secure base: Clinical applications of attachment theory.* London: Routledge.

Burns, D. D., & Nolen-Hoeksema, S. (1992). Therapeutic empathy and recovery from depression in cognitive-behavioral therapy: A structural equation model. *Journal of Consulting and Clinical Psychology, 60,* 441-449.

Burstein, J. W., & Carkhuff, R. R. (1968). Objective therapist and client ratings of therapist-offered facilitative conditions of moderate to low functioning therapists. *Journal of Clinical Psychology, 24, 240-241.*

Busch, A. M., Kanter, J. W., Landes, S. J., & Kohlenberg, R. J. (2006). Sudden gains and outcome: A broader temporal analysis of cognitive therapy for depression. *Behavior Therapy, 17,* 61-68.

Busseri, M. A., & Tyler, J. D. (2003). Interchangeability of the Working Alliance Inventory and Working Alliance, Short Form. *Psychological Assessment, 15,* 193-197.

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

Carkhuff, R. R., Kratochvil, D., & Friel, T. (1968). Effects of professional training: Communication and discrimination of communication. *Journal of Clinical Psychology, 15,* 68-74.

Carmin, C. N., & Dowd, E. T. (1988). Paradigms in cognitive psychotherapy. In W. Dryden & P. Trower (Eds.), *Developments in cognitive psychotherapy* (pp. 1-22). London: Sage Productions.

Carroll, K. M., Nich, C., & Ball, S. A. (2005). Practice makes progress? Homework assignments and outcome in treatment of cocaine dependency. *Journal of Consulting and Clinical Psychology, 73,* 749-755.

Carroll, K. M., Nich, C., & Roundsaville, B. J. (1997). Contribution of the therapeutic alliance to outcome in active versus control psychotherapies. *Journal of Consulting and Clinical Psychology, 65,* 510-514.

Cashdam, S. (1973). *Interactional psychotherapy: Stages and strategies in behavioral change.* New York: Grune & Stratton.

Castonguay, L. G., Constantino, M. J., & Holtforth, M. G. (2006). The working alliance: Where are we and where should we go? *Psychotherapy: Theory, Research, Practice, Training, 43,* 271-279.

Castonguay, L. G., Goldfried, M. R., Wiser, S., Raue, P. J., & Hayes, A. M. (1996). Predicting the effect of cognitive therapy for depression ; A study of unique and common factors. *Journal of Consulting and Clinical Psychology, 64,* 497-504.

Cecero, J. J., Fenton, L. R., Nich, C., Frankforter, T. L., & Carroll, K. M. (2001). Focus on therapeutic alliance: The psychometric properties of six measures across three treatment. *Psychotherapy, 38,* 1-11.

Chambless, D. L., & Hollon, S. (1998). Defining empirically supported therapies. *Journal of Consulting and Clinical Psychology, 66,* 7-18.

Chen, E. C., & Bernstein, B. L. (2000). Relations of complementarity and supervisory issues to supervisory working alliance: A comparative analysis of two cases. *Journal of Counseling Psychology, 47.* 485-497.

Cicchetti, D. V. (1994). Guidelines, criteria, and rules for thumb for evaluating normed and standardized assessment instruments in psychology. *Psychological Assessment, 6,* 284-290.

Cicchetti, D. V., & Sparrow, S. S. (1981). Developing criteria for establishing interrater reliability of specific items: Applications to assessment of adaptive behavior. *American Journal of Mental Deficiency, 84,* 127-137.

Clark, D. A. (1995). Perceived limitations of standard cognitive therapy: A consideration of efforts to revise Beck's theory and therapy. *Journal of Cognitive Psychotherapy: An International Quarterly, 9,* 153-172.

Clark, D. A., & Beck, A. T. (1999). *Scientific foundations of cognitive theory and therapy of depression.* New York: John Wiley & Sons.

Colli, A. & Lingiardi, V. (2009). The Collaborative Interactions Scale: A new transcript-based method for the assessment of therapeutic alliance ruptures and resolutions in psychotherapy. *Psychotherapy Research, 19,* 718-734.

Collins, L.M., & Graham, J. W. (2002). The effect of timing and spacing of observations in longitudinal studies of tobacco and other drug use: temporal design considerations. *Drug and Alcohol Dependence, 68,* 585-596.

Connolly Gibbons, M. B., Crits-Christoph, P., De La Cruz, C., Barber, J., Siqueland, L., & Gladis, M. (2003). Pretreatment expectations, interpersonal functioning, and symptoms in the prediction of the therapeutic alliance across supportive-expressive psychotherapy and cognitive therapy. *Psychotherapy Research, 13,* 59-76.

Constantino, M. J., Arnow, B. A., Blasey, C., & Agras, W. S. (2005). The association between patient characteristics and therapeutic alliance in cognitive-behavioral and interpersonal therapy for bulimia nervosa. *Journal of Consulting and Clinical Psychology, 73,* 203-211.

Constantino, M. J., Castonguay, L. G., & Schut, A. J. (2002). The working alliance: A flagship for the 'scientist-practitioner" model in psychotherapy. In G. S. Tryon (Ed.), *Counseling based on process research: Applying what we know* (pp. 81-131). Boston: Allyn & Bacon.

Cook, J. E., & Doyle, C. (2002). Working alliance in online therapy as compared to face-to-face therapy: Preliminary results. *CyberPsychology & Behavior, 5,* 95-105.

Coon , D. W., Rabinowitz, Y. G., Thompson, L. W., & Gallagher-Thompson, D. (2005). Older adults. In N. Kazantzis, F. P. Deane, K. R. Ronan, & L. L'Abate (Eds.), *Using homework assignments in cognitive behavior therapy* (pp. 117-152). New York: Routledge.

Cooper, M. (2003). *Existential therapies.* London: Sage Publications.

Cooper, J. O., Heron, T.E., & Heward, W. L. (1987). *Applied behavior analysis.* Columbus: Merrill Publishing Co.

Corey, G. (2005). *Theory and practice of counseling and psychotherapy* (7th ed.). Pacific Grove, CA: Brooks/Cole.

Craighead, W. E., Sheets, E. S., Brosse, A. L., & Ilardi, S. S. (2007). Psychosocial treatments for major depressive disorder. In P. E. Nathan & J. M. Gorman (Eds.), *A guide to treatments that work* (pp. 289-307). Oxford: Oxford University Press.

Critchfield, K. L., & Benjamin, L. S. (2006). Integration of therapeutic factors in personality disorders. In L. G. Castonguay & L.E. Beutler (Eds.), *Principles of therapeutic change that work* (pp.253-271). Oxford: Oxford University Press.

Crits-Christoph, P., Connolly Gibbons, M. B., & Hearon, B. (2006). Does the alliance cause good outcome? Recommendations for future research of the alliance. *Psychotherapy: Theory, Research, Practice, Training, 43,* 280-285.

Cuijpers, P., van Straten, A., Andersson, G., & van Oppen, P. (2008). Psychotherapy for depression in adults: A meta-analysis of comparative outcome studies. *Journal of Consulting and Clinical Psychology, 76,* 909-922.

Cuijpers, P., van Straten, A., Schuurmans, J., van Oppen, P., Hollon, S. D. & Andersson G, (2010). Psychotherapy for major depression and dysthymia: A meta-analysis. *Clinical Psychology Review, 30,* 51-62.

Dattilio, F. M. (2001). Cognitive-behavioral family therapy: Commons myths and misconceptions. Con*temporary Family Therapy*, 23, 3-18.

Dattilio, F. M. (2004). *Cognitive behavioral strategies with couples and families.* Seminar presented at New Zealand Psychological Society and School of Psychology, Massey University. August 6th, 2004. Auckland, New Zealand.

Dattilio, F. M., Freeman, A., & Blue, J. (1998). The therapeutic relationship. In A. S. Belleck & M. Hersen (Eds.), *Comprehensive clinical psychology* (volume 6, pp. 229-249). New York: Pergamon Press.

DeCotiis, T. A. (1977). An analysis of the external validity and applied relevance of three rating formats. *Organizational Behavior and Human Performance, 19,* 247-266.

de Roten, Y., Darwish, J., Stern, D. J., Fivaz-Depeuringe, E., Corboz-Warney, A. (1999). Nonverbal communication and alliance in therapy: The body formation coding system. *Journal of Clinical Psychology, 55*, 425-438.

DeRubeis, R. J., Brotman, M. A., & Gibbons C. J. (2005). A conceptual and methodological analysis of the nonspecifics argument. *Clinical Psychology: Science and Practice, 12,* 174-183.

DeRubeis, R. J., & Crits-Cristoph, P. (1998). Empirically-supported individual and group psychological treatments for adult mental disorders. *Journal of Consulting and Clinical Psychology, 66,* 37-52.

DeRubeis, R. J., & Feeley, M. (1990). Determinants of change in cognitive therapy for depression. *Cognitive Therapy and Research, 14,* 469-482.

DeRubeis, R. J., Gelfand, L. A., Tang, Z., & Simons, A. D. (1999). Medications versus cognitive behavior therapy for severely depressed outpatients: Mega-analysis of four randomized comparisons. *American Journal of Psychiatry, 156,* 1007-1013.

Dies, R. R. (2003). Group psychotherapies. In A. Gurman & S. Messer (Eds.), *Essential psychotherapies: Theory and Practice* (2nd ed, pp.515-550). New York: Guilford Press.

Dimidjian, S., Hollon, S. D., Dobson, K. S., Schmaling, K. B., Kohlenberg, R. J., Addis, M. E., Gallop, R., McGlinchey, J. B., Markley, D. K., Gollan, J. K., Atkins, D. K., Dunner, D. l., & Jacobson, N. S. (2006). Randomized trial of behavioral activation, cognitive therapy, and antidepressant medication in the acute treatment of adults with major depression. *Journal of Consulting and Clinical Psychology, 74,* 658-670.

Dinger, U., Strack, M., Leichsenring, F., & Schauenburg, H. (2007). Influences of patients' and therapists' interpersonal problems and therapeutic alliance on outcome in psychotherapy. *Psychotherapy Research, 17,* 148-159.

Dobson, K. (1989). A meta-analysis of the efficacy of cognitive therapy for depression. *Journal of Consulting and Clinical Psychology, 57,* 414-419.

Dobson K., & Dozois, D. (2001). Historical and philosophical bases of the cognitive-behavioral therapies. In K. Dobson (Ed.), *Handbook of cognitive-behavioral therapies* (pp. 3-39). New York: Guilford Press.

Dobson, K. S., Shaw, B. F., & Vallis, T. M. (1985). Reliability of a measure of the quality of cognitive therapy. *British Journal of Clinical Psychology. 24,* 295-300.

Dozois, D. J. A., Dobson K. S., & Ahnberg, J .L. (1998). A psychometric evaluation of the Beck Depression Inventory-II. *Psychological Assessment,* 10, 83-89.

Dryden, W., & Ellis, A. (2001). Rational emotive behavior therapy. In K. Dobson (Ed.), *Handbook of cognitive-behavioral therapies* (pp. 295-348). New York: Guilford Press.

Dryden, W., & Trower, P. (1988). *Developments in cognitive psychotherapy.* London: Sage Productions.

Dunn, H., Morrison, A. P., & Bentall, R. P. (2002). Patients experiences of homework tasks in cognitive behavioral therapy for psychosis: a qualitative analysis. *Clinical Psychology and Psychotherapy, 9,* 361-369.

Dunn, H., Morrison, A. P., & Bentall, R. P. (2006). The relationship between patient suitability, therapeutic alliance, homework compliance and outcome in cognitive therapy for psychosis. *Clinical Psychology and Psychotherapy, 13,* 145-152.

Eaton, T. T., Abeles, N., & Gutfreund, M. J., (1988). Therapeutic alliance and outcome: Impact of treatment length and pretreatment symptomology. *Psychotherapy, 25,* 536-542.

Ekman, P., & Friesen, W. V. (1968). Nonverbal behavior in psychotherapy research. In J.M. Shlien (Ed), *Research in psychotherapy: proceedings of the third conference, Chicago, Illinois, May 31-June 4, 1966* (pp. 179-216). Washington, DC: American psychological Association.

Elkin, I. (1999). A major dilemma in psychotherapy outcome study: Disentangling therapist form therapies. *Clinical Psychology: Science and Practice, 6,* 10-32.

Elkin, I., Shea, M. T., Watkins, J. T., Imber, S. D., Sotsky, S. M., Collins, J. F., Glass, D. R., Pilkonis, P. A., Leber, W. R., Doherty, J. P., Fiester, S. J., & Parloff, M. B. (1989). National Institute of Mental Health Treatment of Depression Collaborative Research Program. *Archives of General Psychiatry, 46*, 971-982.

Elliot, R., Clark, C., Kemeny, V., Wexler, M., Mack, C., & Brinkerhoff, J., (1990). The impact of experiential therapy on depression: The first ten cases. In G. Lietaer, J. Rombauts, & R. Van Balen (Eds.)*, Client-centered therapy and experiential therapy In the nineties.* Belgium: Leuven University Press.

Elliott, R., & Greenberg, L. S. (2002). Process-experiential psychotherapy. In J. J. C. Cain & J. Seeman (Eds.), *Humanistic psychotherapies: Handbook of research and practice* (pp. 279-306). Washington, DC: American Psychological Association.

Elliot, R., & Greenberg, L. S. (1995). Experiential therapy in practice: The process-experiential approach. In B.Bongar & L.E. Beulter (Eds), *Comprehensive textbook of psychotherapy* (pp.123-139). New York: Oxford University Press.

Ellis, A. (1962). *Reason and emotion in psychotherapy.* Secuacus, NJ: Lyle Stuart.

Ellis, A. (1973). *Humanistic psychotherapy: The rational-emotive approach.* New York: McGraw-Hill.

Ellis, A. (1996). *Better, deeper, and more enduring brief therapy: the rational emotive behavior therapy approach.* New York: Brunner/Mazel, Inc.

Ellis A., & Dryden, W. (1997). *The practice of rational-emotive behaviour therapy* (2nd ed.). New York; Springer Pub Co.

Elvins, R. & Green, J. (2008). The conceptulization and measurement of therapeutic alliance: An empirical review. *Clinical Psychology Review, 28*, 1167-1187.

Erdman, H. P., Klien, M. H., & Greist, J. H., Direct patient computer interviewing. *Journal of Consulting and Clinical Psychology, 53,* 760-773.

Fava, G. A., Ruini, C., Rafanelli, C., Finos, L., Conti, S. & Grandi, S. (2004). Six-year outcome of cognitive behavior therapy for prevention of recurrent depression. *American Journal of Psychiatry, 161,* 1872-1876.

Feeley, M., DeRubeis, R. J., & Gelfand, L. A. (1999). The temporal relation of adherence to symptom change in cognitive therapy for depression. *Journal of Consulting and Clinical Psychology, 67,* 578-582.

Fehm, L., & Kazantzis, N., (2004). Attitudes and the use of homework assignments in therapy: A survey of German psychotherapists. *Clinical Psychology & Psychotherapy 11,* 332-343.

Feldman, K. A. (1978). Course characteristics and college student's ratings of their teachers: What we know and what we don't. *Research in Higher Education, 9,* 199-242.

Fenton, L. R., Cecero, J. J., Nich, C., Fankforter, T. L., & Carroll, K. M. (2001). Perspective is everything: The predictive validity of six working alliance instruments. *Journal of Psychotherapy Practice and Research, 10,* 262-268.

Fleiss, J. L. (1973). *Statistical methods for rates and proportions.* New York: John Wiley & Sons.

Fleiss, J. L. (1981). *Statistical methods for rates and proportions (2nd ed.).* New York: John Wiley & Sons.

Foa, E. B., & Rothbaum, B. O. (1998). *Treating the trauma of rape: Cognitive Behavioral therapy for PTSD.* New York: Guilford Press.

Ford, J. D. (1978). Therapeutic relationship in behavior therapy: An empirical analysis. *Journal of Consulting and Clinical Psychology, 46,* 1302-1314.

Freeberg, N. E. (1969). Relevance of rater-ratee acquaintance in the validity and reliability of ratings. *Journal of Applied Psychology, 61,* 518-524.

Freud, S. (1910). *The future prospects of psychoanalytic therapy. Standard Edition,* 11, 139-151. London: Hogarth Press, 1957.

Friedman, M. A., & Whisman, M. A. (2004). Implicit cognition and the maintenance and treatment of major depression. *Cognitive and Behavioral Practice, 11,* 168-177.

Gaffan, E. A., Tsaousis, I., & Kemp-Wheeler, S. M. (1995). Researcher allegiance and meta-analysis: The case of cognitive therapy for depression. *Journal of Consulting and Clinical Psychology, 63,* 966-980.

Gaston, L. (1990). The concept of the alliance and its role in psychotherapy: Theoretical and empirical considerations. *Psychotherapy Research, Practice, Training, 27,* 143-153.

Gaston, L., Goldfried, M. R., Greenberg, L. S., Horvath, A. O., Raue, P. J., & Watson, J. (1995). The therapeutic alliance in Psychodynamic, cognitive-behavioral, and experiential therapies. *Journal of Psychotherapy Integration, 5,* 1-26.

Gaston L., & Marmar, C .R. (1994). The California Psychotherapy Alliance Scales. In A. O. Horvath & L. S. Greenberg (Eds.), *The working alliance: Theory, research and practice* (pp. 1-9). New York: John Wiley & Sons.

Gaston, L., Marmar, C. R., Gallagher, D., & Thompson, L. W. (1991). Alliance prediction of outcome beyond in-treatment symptomatic change as psychotherapy processes. *Psychotherapy Research, 1,* 104-113.

Gaston, L., Marmar, C. R., Thompson, L. W., & Gallagher, D. (1988). Relation of patient pretreatment characteristics to the therapeutic alliance in diverse psychotherapies. *Journal of Consulting and Clinical Psychology, 56,* 483-489**.**

Gaston, L., & Ring, J. M. (1992). Preliminary results on the inventory of therapeutic strategies. *Journal of Psychotherapy Practice and Research, 1,* 135-146.

Gaston, L., Thompson, L., Gallagher, D., Cournyer, L., & Gagnon, R. (1998). Alliance, technique, and their interactions in predicting outcome of behavioral, cognitive and brief dynamic therapy. *Psychotherapy Research, 8,* 190-209.

Gelso, C. J., & Carter J. A. (1985). The relationship in counseling and psychotherapy: Components, consequences, and theoretical antecedents. *The Counseling Psychologist, 13,* 155-244.

Gelso, C. J., & Carter J. A. (1994). Components of the psychotherapy relationship: Their interaction and unfolding during treatment. *Journal of Counseling Psychology, 41,* 196-306.

Gilbert, P., & Leahy, R. L. (2007). Introduction and overview: Basic issues in the therapeutic relationship. In P. Gilbert and R. L. Leahy (Eds.), *The therapeutic relationship in the cognitive behavioral psychotherapies* (pp. 3-23). East Sussex: Routledge.

Giovazolias, T. (2004). The therapeutic relationship in cognitive-behavioural therapy. *Counselling Psychology Review, 19,* 14-20.

Gladding, S. T. (2007). *Counseling: A comprehensive profession* (5th ed.). Upper Saddle River, New Jersey: Pearson Prentice Hall.

Glaser, B. G. & Strauss, A. (1967). *The discovey of grounded theory: Strategies for qualitative research.* Chicago IL: Aldine.

Golden, B. R., & Robbins, S. B. (1990). The working alliance within time-limited therapy: A case analysis. *Professional Psychology: Research and Practice, 21,* 476-481.

Goldfried, M. R. (2003). Cognitive-behavior therapy: Reflections of the evolution of a therapeutic orientation. *Cognitive Therapy and Research, 27,* 53-69.

Goldfried, M. R. (2007). What has psychotherapy inherited from Carl Roger's? *Psychotherapy: Theory. Research, Practice, Training, 44*, 249-252.

Goldfried, M. R., & Davison, G. C. (1976). *Clinical behavior therapy.* New York: Holt, Rinehart and Winston.

Greenberg, L. S. & Safran, J. D. (1984).Integrating effect and cognition: A perspective on the process of therapeutic change. *Cognitive Therapy and Research, 8*, 591-598.

Greenberg, L. S. & Safran, J. D. (1987). *Emotion in psychotherapy.* New York: Guilford Press

Greenson, R. R. (1965). The working alliance and the transference neurosis. *The Psychoanalytic Quarterly, 34,* 15-181.

Haggard, E. A. (1958). *Intraclass correlation and the analysis of variance.* New York: The Dryden Press, Inc.

Hall, S. F., Groome, P. A., Streiner, D. L. & Rochon, P. A., (2006). Interrater reliability and measurements of comorbid illness. *Journal of Clinical Epidemiology, 59*, 926-933.

Hambleton, R. K., & Powell, S. (1983). A framework for viewing the process of standard setting. *Evaluation and Health professionals, 6,* 3-24.

Hamilton, M. (1960). A rating scale for depression. *Journal of Neurology, Neurosurgery,and Psychiatry, 23, 56*-62.

Hanson, W. E., Curry, K. T., & Bandalos, D. L. (2002). Reliability generalization of the Working Alliance Inventory Scale scores. *Educational and Psychological Measurement. 62,* 659-673

Hardy, G. E., Cahill, J., Shapiro, D. A., Barkham, M., Rees, A., Macaskill, N. (2001). Client interpersonal and cognitive styles as predictors of response to time-limited cognitive therapy for depression. *Journal of Consulting & Clinical Psychology, 69,* 841-845.

Hartley, D., & Strupp, H. (1983). The therapeutic alliance: Its relationship to outcome in brief psychotherapy. In J. Masling (Ed.) Empirical studies of psychoanalytic theories (Vol. 1, pp. 1-27). Hillsdale, NJ: Erlbaum.

Hatcher, R.L. (2006). How a return to theory could help alliance research. *Psychotherapy: Theory, Research, Practice, Training,* 43, 292-299.

Hatcher, R. L., & Barends A., (1996). Patients' view of the alliance in psychotherapy: Exploratory factor analysis of three alliance measures. *Journal of Consulting and Clinical Psychology, 64,* 1326-1336.

Hatcher, R. L., Barends, A., Hansell, J., & Gutfreund, M. J. (1995). Patients' and therapists' shared unique views of the therapeutic alliance: An investigation using confirmatory factor analysis in a nested design. *Journal of Consulting and Clinical Psychology, 63,* 636-643.

Hatcher, R. L., & Gillaspy, A. (2006). Development and validation of a revised short version of the Working Alliance Inventory. *Psychotherapy Research, 16,* 12-25.

Hay, P., Katsikitis, M., Begg, J., Da Costa, J., & Blumenfield, N. (2003). A two-year follow-up study and prospective evaluation of the DSM-IV Axis V. *Psychiatric Services, 54,* 1028-1030.

Hayes, A. M., Beevers, C., Feldman, G., Laurenceau, J., & Perlman, C. A. (2005). Avoidance and processing as predictors of symptom change and positive growth in an integrative therapy for depression. *International Journal of Behavioral Medicine, 12,* 111-122.

Hayes, A. M., Laurenceau, J., Feldman, G., Strauss, J. L., & Cardaciotto, L. (2007). Change is not always linear: the study of nonlinear and discontinuous patterns of Change in psychotherapy. *Clinical Psychology Review, 27,* 715-723.

Hayes, S. C. (1998). Single case experimental design and empirical clinical practice. In A. E. Kazdin (Ed.), *Methodological Issues & strategies in clinical research* (2nd ed.) (pp. 419-449). Washington, DC: American Psychological Association.

Heneman, H. G., Schwab, D. P., Huett, D. L., & Ford, J. J. (1974). Interviewer validity as a function of interview structure, biographical data, and interview order. *Journal of Applied Psychology, 60,* 748-753.

Henry, W. P., & Strupp, H. H. (1994). The therapeutic alliance as interpersonal process. In A. O. Horvath & L. S. Greenberg (Eds.), *The working alliance: theory research and practice* (pp. 51–84). New York: John Wiley & Sons.

Hersen, M., & Barlow, D. H. (1976). *Single-case experimental design: Strategies for studying behavior change.* New York: Pergamon Press.

Hersoug, A.G. (2001). Quality of working alliance in psychotherapy: Therapist variables and patient/therapist similarity as predictors. *Journal of Psychotherapy Practice and Research, 10,* 205-216.

Hersoug, A. G., Monsen, J. T., Havik, O. E., & Hoglend, P. (2002). Quality of early working alliance in psychotherapy: Diagnoses, relationship and intrapsychic variables as predictors. *Psychotherapy and Psychosomatics, 71,* 18-27.

Hewitt, J., & Coffey, M. (2005). Therapeutic relationships with people with schizophrenia: Literature review. *Journal of Advanced Nursing, 5,* 561-570.

Hill, C. E., Carter, J. A., & O'Farrell, M. K. (1983). A case study of the process and outcome of time limited counseling. *Journal of Counseling Psychology, 30,* 3-18.

Hill, C. E., & Lambert, M. J. (2004). Methodological issues in studying psychotherapy processes and outcomes. In M.J. Lambert (Ed.), *Bergin and Garfield's Handbook of psychotherapies and behavior change* (5th ed., p. 84-135). New York: John Wiley & Sons.

Hill, C. E., Nutt, E., & Jackson, S. (1994). Trends in psychotherapy process research: Samples, measures, researchers, and classic publications. *Journal of Counseling Psychology, 41,* 364-377.

Hill, C. E. & O'Grady, K. E. (1985). List of therapist intentions illustrated in a case study and with therapists of varying theoretical orientations. *Journal of Counseling Psychology, 32,* 3-22.

Hill, C. E., O'Grady, K. E., Elkin, I. (1992). Applying the Collaborative Study Psychotherapy Rating Scale to rate therapist adherence in cognitive-behavior therapy, interpersonal therapy, and clinical management *Journal of Consulting and Clinical Psychology, 60,* 73-79.

Hill, C.E., & Stephany, A. (1990). The relationship of nonverbal behaviors to client reactions. *Journal of Counseling Psychology, 37,* 22-26.

Hillard, R. B., Henry, W. P., & Strupp, H. H. (2000). An interpersonal model of psychotherapy: Linking patient and therapist developmental history, therapeutic process, and types of outcome. *Journal of Consulting and Clinical Psychology, 68,* 125-133.

Hilsenroth, M. J., Ackerman, S.J., Blagys, M D., Baumann, B.D., Baity, M.R., Smith, S.R., Price, J.L., Smith, C.L., Heindselman, T.L., Mount, M.K.,

Holdwick, D.J Jr. (2000). Reliability and validity of DSM-IV axis V. *American Journal of Psychiatry, 157,* 1858-1863.

Hilsenroth, M. J., & Cromer, T. D. (2007). Clinician interventions related to alliance during the initial interview and psychological assessment. *Psychotherapy: Theory, Research, Practice, Training, 43,* 205-218.

Hirschfeld, R. M. A. (2001). The comorbidity of major depression and anxiety disorders: Recognition and management in primary care. *Primary Care Companion, Journal of Clinical Psychiatry,* 3, 244-254.

Hollan, S. D., & Beck, A. T. (1979). Cognitive therapy of depression. In P. C. Kendall & S. D. Hollon (Eds.), *Cognitive-behavioral interventions* (pp. 153-204). New York: Academic Press.

Hollan, S. D., & Beck, A. T. (2004). Cognitive and cognitive behavioral therapies. In M. J. Lambert (Ed.), *Bergin and Garfield's Handbook of psychotherapies and behavior change* (5th ed., p. 447-492). New York: John Wiley & Sons.

Hollon, S. D., DeRubeis, R. J., Evans, M. D., Wiemer, M. J., Garvey, M. J., Grove, & W. M., Tuason, V. B. (1992). Cognitive therapy and pharmoccotherapy for depression: Singly and in combination. *Archives of General Psychiatry. 49,* 774-781.

Hollon, S. D., DeRubeis, R. J., Shelton, R. C., Amsterdam, J. D., Salomon, R. M., O'Reardon, J. P., Lovett, M. L., Young, P. R., Haman, K. L., Freeman, B. B., Gallop, R. (2005). Prevention of relapse following cognitive therapy vs medications in moderate to severe depression. *Archives of General Psychiatry. 62,* 417-422.

Holzbach, R. (1978). Rater bias in performance ratings: Superior, self, and peer ratings. *Journal of Applied Psychology, 63,* 579-588.

Horvath, A. O. (1981). *An exploratory study of the working alliance; Its measurement and relationship to outcome.* Unpublished doctoral thesis. University of British Columbia, Burnaby, Canada.

Horvath, A. O. (1994a). Empirical validation of Bordin's pantheoretical model of the alliance. In A. O. Horvath & L. S. Greenberg (Eds.), *The working alliance: Theory, research and practice* (pp. 109-128). New York: John Wiley & Sons.

Horvath, A. O. (1994b). Research on the alliance. In A. O. Horvath & L. S. Greenberg (Eds.), *The working alliance: Theory, research and practice* (pp. 259-286). New York: John Wiley & Sons.

Horvath, A. O. (2000). The therapeutic relationship: From transference to alliance. *Journal of Clinical Psychology, 56,* 163-173.

Horvath, A. O. (2006). *Update on the short form of the Working Alliance Inventory*. (Google). http://www.educ.sfu.ca/alliance/alianceA

Horvath, A. O., & Greenberg, L. S. (1989). Development and validation of the Working Alliance Inventory. *Journal of Counseling Psychology, 36,* 223-233.

Horvath, A. O., & Greenberg, L. S. (1994).Introduction. In A. O. Horvath & L. S. Greenberg (Eds.), *The working alliance: Theory, research and practice* (pp. 1-9). New York: John Wiley & Sons

Horvath, A. O., & Luborsky, L. (1993). The role of the therapeutic alliance in psychotherapy. . *Journal of Consulting and Clinical Psychology, 61,* 561-573.

Horvath, A. O., & Marx, R. W. (1990). The development and decay of the working alliance during time-limited counselling. *Canadian Journal of Counselling, 24,* 240-260.

Horvath, A. O., & Symonds, B. D. (1991). Relation between working alliance and outcome in psychotherapy; A meta-analysis. *Journal of Counseling Psychology, 38,* 139-149.

Howgego, I. M., Yellowlees, P., Owen, C., Meldrum, L., & Dark, F. (2003). The therapeutic alliance: The key to effective patient outcome? A descriptive review of the evidence in community mental health case management. *Australian & New Zealand Journal of Psychiatry, 37,* 169-183.

Ilardi, S. S., & Craighead, W. E. (1994). The role of nonspecific factors in cognitive-behavioral therapy for depression. *Clinical Psychology: Science and Practice, 1,* 138-156.

Imel, Z. E., Malterer, M. B., McKay, K. M., & Wampold, B. E. (2008). A meta-analysis of psychotherapy and medication in unipolar depression and dysthymia. *Journal of Affective Disorders, 110,* 197-206.

Ivancevich, J. M. (1979. Longitudinal study of the effects of rater training on psychometric error in ratings. *Journal of Applied psychology, 64,* 502-508.

Jacobson, N. S., Dobson, K. S., Traux, P. A., (1996). A component analysis of cognitive-behavioral treatment for depression. *Journal of Consulting and Clinical Psychology, 64,* 295-304

Jaeger, R. M, & Busch, J.C. (1984). *The effects of a Delphi modification of the Anoff-Jaeger standard setting procedure on standards recommended for the National Teacher Examination.* Paper presented at the annual meeting of the National Council on Measurement in Education, New Orleans, LA.

Jones, E. E. (1978) Effects of race on psychotherapy process and outcome: An exploratory Investigation *Psychotherap: Theory, Research and Practice, 15,* 226-236.

Jones, E. E. & Zoppel, C. L. (1982). Impact of client and therapist gender on psychotherapy process and outcome. *Journal of Consulting and Clinical Psychology, 50,* 259-272.

Joyce, A. S., Ogrodniezuk, J. S. Piper, W. E., & McCallum, M. (2003). The alliance as a mediator of expectancy effects in short-term individual therapy. *Journal of Consulting and Clinical Psychology, 38,* 139-149.

Joyce, A. S., & Piper, W. E. (1998). Expectancy, the therapeutic alliance, and treatment outcome in short-term individual psychotherapy. *Journal of Psychotherapy Practice and Research, 7,* 236-248.

Kaplan, R. M., & Saccuzzo, D. P. (2001). *Psychological testing: Principles applications, and issues* (5th ed.), Belmont, CA: Wadswoth/Thompson.

Kazantzis, N., & Deane, F. P. (1998). Theoretical orientations of New Zealand Psychologists: An international comparison. *Journal of Psychotherapy Integration, 8,* 97-113.

Kazantzis, N., Lampropoulos, G. K., & Deane, F. P. (2005). A national survey of practicing psychologists use and attitudes toward homework in psychotherapy. *Journal of Consulting and Clinical Psychology, 73,* 742-748.

Kazantzis, N., MacEwan, J., & Dattilio, F. M. (2005). A guiding model for practice. In N. Kazantzis, F. P. Deane, K. R. Ronan, & L. L'Abate (Eds.), *Using homework assignments in cognitive behavior therapy* (pp. 357-404). New York: Routledge.

Kazantzis, N., Wedge, P., & Dobson, K. S. (2005). *The Homework Adherence and Competence Scale (HAACS).*

Kazdin, A. E. (1977). Artifact, bias, and complexity of assessment: The ABCs of reliability. *Journal of Applied behavior Analysis, 10,* 141-150.

Kazdin, A. E. (1998). Drawing valid inferences from case studies. In A. E. Kazdin (Ed.), *Methodological Issues & strategies in clinical research* (2nd ed., pp. 403-417). Washington, DC: American Psychological Association.

Kazdin, A. E. (2001).progression of therapy research and clinical application of Treatment require better understanding of the change process. *Clinical Psychology: Science and Practice, 8*, 143-151.

Kazdin, A. E. (2003). *Research design in clinical psychology.* Boston: Allyn & Bacon.

Kazdin, A. E. (2005). Treatment outcomes, common factors and continued neglect of mechanisms of change. *Clinical Psychology: Science and Practice, 12,* 184-188.

Kazdin, A. E. (2006). Mechanisms of change in psychotherapy: Advances, breakthroughs and cutting-edge research (do not yet exist). In R. R. Bootzin & P. E. Knight (Eds.), *Strenghtening research methodology: Psychological measurement and evaluation,* (pp. 77-101). Washington, DC: American Psychological Association.

Kazdin, A. E. (2010). Single-case research designs: methods for clinical and applied Settings (2nd ed.). Oxford: Oxford University Press.

Keaveany, T. J., & McGann, A. F. (1975). A comparison of behavioral expectation scales and graphic rating scales *Journal of Applied Psychology, 60,* 695-703.

Keen, A. J. A., & Freeston, M. H. (2008). Assessing competence in cognitive-behavioural therapy. *The British Journal of Psychiatry, 193,* 60-64.

Kelly, G. (1955). *The Psychology of Personal Constructs.* New York: Norton.

Kennedy, N., Abbott, R., & Paykel, E. S. (2003). Remission and recurrence of depression in the maintenance era: Long- term outcome in Cambridge. *Psychological Medicine, 33,* 827-838.

Khan-Bourne, N., & Brown, R. G. (2003). Cognitive behaviour therapy for the treatment of depression in individuals with brain injury. *Neuropsychological Rehabilitation, 13,* 89-107.

Kiesler, D.J., (1973). *The process of psychotherapy: Empirical foundations and systems of analysis.* Chicago: Aldine Pub. Co.

Kim, I. (1995). *Examining ways to improve quality of ratings.* Unpublished doctoral thesis,The Ohio State University, Ohio, U.S.A.

King, L. M., Hunter, J. E., & Schmidt, F. L. (1980). Halo in a multidimensional forced-choice evaluation scale. *Journal of Applied Psychology, 65,* 507-516.

Kirschenbaum, H., & Jordan, A. (2005). The current status of Carl Rogers and the person- centered approach. *Psychotherapy: Theory, Research, Practice and Training, 42,* 37-51.

Kivlighan, D. M., Patton, M. J., & Foote, D. (1998). Moderating effects of client attachment on the counselor experience-working alliance relationship. *Journal of Counseling Psychology, 45,* 274-278.

Kivilighan, D. M., & Schmitz, P. J. (1992). Counselor technical activity in cases with improving working alliances and continuing-poor working alliances. *Journal of Counseling Psychology, 39,* 32-38.

Kivlighan, D. M., & Shaughnessy, P. (1995). An analysis of the development of the working alliance using hierarchical linear modeling. *Journal of Counseling Psychology, 42,* 338-349.

Klerman, G., & Weissman, M. (1992). The course, morbidity, and costs of depression. *Archives of General Psychiatry, 49,* 831-834.

Klien, M. J., Mathieu-Coughlan, P., & Kiesler, D. J. (1986). The Experiencing Scales. In L. S. Greenberg & W. M. Pinsof (Eds.), The psychotherapeutic process: A research handbook (pp. 21-71). New York: Guilford.

Klien, D. N., Schwartz, J. E., Castonguay, L. G., Arnow, B., Blalock, J. A., Manber, R., Markowitz, J. C., Riso, L. P., Rothbaum, B., McCollough., J. Thase. M. E. Borian, F. E., Miller, I. W., & Keller, M.B. (2003). Therapeutic alliance in depression treatment: Controlling for prior change and patient characteristics. *Journal of Consulting and Clinical Psychology, 71,* 997-1006.

Knaevelsrud, C., & Maercker, A. (2006). Does the quality of the working alliance predict treatment outcome in online psychotherapy for traumatized patients. *Journal of Medical Internet Research, 8,* Article 31.

Knight R.G. (1984). Some general population norms for the short form Beck Depression Inventory. *Journal of Clinical Psychology, 40,* 751-753.

Kobak, K. A., Engelhardt, N., Williams, J. B. W., & Lipsitz, J. D. (2008). Rater training in multicenter clinical trials: Issues and recommendations. *Journal of Clinical Psychopharmacology, 24,* 113-115.

Kokotovic, A. M., & Tracey, T. J. (1990). Working alliance in the early phase of counseling. *Journal of Counseling Psychology, 37,* 16-21.

Komiti, A. A., Jackson, H. J., Judd, F. K., Cockram, A. M., Kyrios, M., Yeatman, R., Murray, G., Hordern, C., Wainwright, K., Allen, N., & Singh, B. (2001). A comparison of the Composite International Diagnostic Interview (CIDI-Auto) with clinical assessment in diagnosing mood and anxiety disorders. *Australian and New Zealand of Psychiatry, 35,* 224-230.

Koong Hean Foo, & Merrick, P. L. (2004). Practitioner characteristics with comparison of counselling and psychotherapy practices between New Zealand and Singapore mental health practitioners. *Proceedings of the Inaugural International Asian Health Conference.*

Korman, A. K. (1977). O*rganization behavior* Englewood Cliffs, N.J: Prentice-Hall.

Kowalik, Z. J., Schiepek, G., Kumpf, K., Roberts, L. E., & Elbert, T. (1997). Psychotherapy as a chaotic process, II: The application of non linear analysis methods on quasi time-series of client-therapist interaction: A nonstationary approach. *Psychotherapy Research, 7,* 179-218.

Kozolowski, S., & Hattrup, K. (1992). A disagreement about within-group agreement: Disentangling issues of consistency versus consensus. *Journal of Applied Psychology, 77,* 161-167.

Krupnick, J. L., Sotsky, S. M., Simmens, S., Moyer, J., Elkin, I., Watkins, J., & Pilkonis, P. A. (1996). The role of the therapeutic alliance in psychotherapy and Pharmacotherapy outcome: Findings in the National Institute of Mental Health Treatment of Depression Collaboration Research Program. *Journal of Consulting and Clinical Psychology, 64,* 532-539.*,*

Kuyken, W. Padesky, C. A. & Dudley, R. (2008). The science and practice of case conceptualization. *Behavioural and Cognitive Psychotherapy, 36,*757-768.

Kuyken, W. Padesky, C. A. & Dudley, R. (2009). *Collaborative case conceptualization: Working effectively with clients in cognitive-behavioral therapy.* New York: Guilford Press.

Lam, D. H., Bright, J., Jones, S., Hayward, P., Schuck, N., Chisolm, D. & Sham, P. (2000) Cognitive therapy for bipolar illness: A pilot study of relapse prevention. *Cognitive Therapy and Research, 24,* 503-520.

Lam, D.H., Watkins, E.R., Hayward, P., Bright, J., Wright, K., Kerr, N., Parr-Davis, G., & Sham, P. (2003). A randomized controlled study of cognitive therapy for relapse prevention for bipolar affective disorder. *Archives of General Psychiatry, 60,* 145-152.

Lambert, M. J., Bergin, A. E., & Garfield, S. L. (2004). Introduction and historical overview. In M. J. Lambert (Ed.), *Bergin and Garfield's Handbook of psychotherapies and behavior change* (5th ed., pp. 3-15). New York: John Wiley & Sons.

Lambert, M., de Julio, S., & Stein, D., (1978). Therapist interpersonal skills. *Psychological Bulletin, 83,* 467-489.

Lambert, M. J., & Ogles, B. (2004). The efficacy and effectiveness of psychotherapy. In M. J. Lambert (Ed.), *Bergin and Garfield's Handbook of psychotherapies and behavior change* (5th ed., p. 139-193). New York: John Wiley & Sons.

Landis, J., & Koch, G. (1977). The measurement of observer agreement for categorical data. *Biometrics, 33,* 159-174.

Langs, R. (2004). *Fundamentals of adaptive psychotherapy and counseling.* United Kingdom: Palgrave MacMillan

Latham, G. P., Wexley, K. N., & Pursell, E. D. (1975). Training managers to minimize rating errors in the observation of behavior. *Journal of Applied Psychology, 60,* 550-555.

Laurenceau, J, Hayes, A. M., & Feldman, G C. (2007). Some methodological and Statistical issues in the study of change processes in psychotherapy. *Clinical Psychology Review, 27,* 682-695.

Leahy, R. L. (2001). *Overcoming resistance in cognitive therapy.* New York: Guilford Press.

Lebow, J., Kelly, J., Knobloch-Fedders, L. M., & Moos, R. (2006). Relationship factors in treating substance use disorders. In L. G. Castonguay & L. E. Beutler (Eds.), *Principles of therapeutic change that work* (pp.293-317). Oxford: Oxford University .

Leitenberg, H., Rosen, J. C., Wolf, J., Vara, L. S. Detzer, M., & Srebnik, D. (1994). Comparison of cognitive-behavior therapy and desipramine in the treatment of bulimia nervosa. *Behavior Research and Therapy, 32,* 37-45.

Levine, J., & Butler, J. (1952). Lecture vs. group discussion in changing behavior. *Journal of Applied Psychology, 36,* 29-33.

Lietaer, G, (1990). The client-centered approach after the Wisconsin project: A personel view on its evolution. . In G. Lietaer, J. Rombauts, & R. Van Balen (Eds.), *Client-centered therapy and experiential therapy in the nineties.* Belgium: Leuven University Press.

Lingiardi, V., Filippucci, L., & Baiocco, R. (2005). Therapeutic alliance evaluation in personality disorders psychotherapy. *Psychotherapy Research, 15,* 45-53.

Llewelyn, S. & Hardy, G. (2001). Process research in understanding and applying psychological therapies. *British Journal of Clinical Psychology, 40,* 1-21.

Loeb, K. L., Wilson, G. T., Labouvie, E., Pratt, E. M., Hayaki, J., Walsh, B. T., Agras, W. S., & Fairburn, C. G. (2005). Therapeutic alliance and treatment adherence in two interventions for bulimia nervosa: A study of process and outcome. *Journal of Consulting and Clinical Psychology, 73,* 1097-1106.

London, M., & Poplawski, J. R. (1976). Effects of information of stereotype development in performance appraisal and interview contexts. *Journal of Applied Psychology, 61,* 199-205.

Luborsky, L. (1994). Therapeutic alliances as predictors of psychotherapy outcomes: Factors explaining predictive success. In A. O. Horvath & L. S. Greenberg (Eds.), *The Theory, research and practice* (pp. 38-50). New York: John Wiley & Sons.

Luborsky, L., Crits-Christoph, P., Alexander, L., Margolis, M., & Cohen M. (1983). Two helping alliance methods for predicting outcomes of psychotherapy: A counting signs vs. a global rating method. *Journal of Nervous and Mental Disease, 171,* 480-491.

Luborsky, L., McLellan, A. T., Woody, G., O,Brien, C. P., & Auerbach, A. (1985). Therapist success and its determinants. *Archives of General Psychiatry, 42*, 602-611.

MacDonald, A. J. (1994). Brief therapy in adult psychiatry. *Journal of Family Therapy, 16,* 415-426.

Mahoney, J. (2000). Strategies of causal inference in small-n analysis. *Sociological Methods Research, 28,* 387-424.

Mallinckrodt, B. (1992). Client's representations of childhood emotional bonds with parents social support, and formation of the working alliance. *Journal of Counseling Psychology, 38,* 401-409.

Mallinckrodt, B., & Nelson, M. L. (1991). Counselor training level and the formation of the psychotherapeutic working alliance. *Journal of Counseling Psychology, 38,* 133-138.

Marmar, G. R., & Gaston, L. (1993). *Manual for The California Psychotherapy Scales CALPAS.* Unpublished Manual, University of California, San Francisco, CA.

Marmar, G. R., Weiss, D. S., & Gaston, L. (1989). Towards the validation of the California Psychotherapy Alliance Rating System. *Journal of Consulting and Clinical Psychology, 1,* 46-52.

Martin, D., Garske, J., & Davis, K. (2000). Relation of the therapeutic alliance with outcome and other variables: A meta-analysis review. *Journal of Consulting and Clinical Psychology, 68,* 438-450.

Marzali, E., Munroe-Blum, H., & McCleary, L. (1999). The effects of the therapeutic alliance on the outcomes of individual and group psychotherapy with borderline personality disorder. *Psychotherapy Research, 9,* 424-436.

May, R., & Yalom, I. (2005). Existential psychotherapy. In R.J. Corsini & D. Wedding (Eds.), *Current psychotherapies* (7th ed, pp. 269-298). Canada: Brooks/Cole, A Division of Thompson Learning, Inc.

McGraw, K. O., & Wong, S. P. (1996). Forming inferences about some intraclass correlation coefficients. *Psychological Methods, 1,* 30-46.

McIntype, R., Smith, D., & Hassett, C. (1984). Accuracy of performance ratings as affected by rater training and perceived purpose of rating. *Journal of Applied Psychology, 69,* 147-156.

McMullin, R. C. (1972). Effects of counselor focusing on client self-experiencing under low attitudinal conditions. *Journal of Clinical Psychology, 19,* 282-285.

Meier, P. S., Barrowclough, C., & Donmall, M. C. (2005). The role of the therapeutic alliance in the treatment of substance misuse: A critical review of the literature. *Addiction, 100,* 304-316.

Mercer, R. C., & Loesch, L. C. (1979). Audio tape ratings: Comments and guidelines: *Psychotherapy, Theory, Research, and Practice, 16,* 79-85.

Mintz, J., Auerbach, A. R., Luborsky, L., & Johnson, M. (1973). Patient's, therapist's and observers' views of psychotherapy: a 'rashomon' experience or a reasonable consensus? *British Journal of Medical Psychology, 46,* 83-89.

Mitchell, S. K. (1979). Inter-observer agreement, reliability, and generalizability of data collected in observational studies. *Psychological Bulletin, 86,* 376-390.

Mitchell, K. M., & Brenson, B. G. (1970). Differential use of confrontation by high-and-low- facilitative therapists. *Journal of Nervous and Mental Disease, 151,* 303-309.

Moras, K., & Hill, C. E. (1991). Rater selection in psychotherapy process research: Observations on the state-of-the-art. *Psychotherapy Research, 1,* 113-123.

Morgan, D. L., & Morgan, R. K. (2003). Single case participant research design: bringing science to managed care. In A.E. Kazdin (Ed.), *Methodological issues and strategies in clinical research (*pp. 635-654). Washington, DC: American Psychological Association.

Moritsch, B. G., & Suter, W. N. (1988). Correlates of halo error in teacher evaluation. *Educational Research Quarterly, 12,* 29-34.

Mueser, K. T., Rosenberg, S. D., Xie, H. Jankowski, M. K., Bolton, E. E., Lu, W., Hamblen, J. L., Rosenberg, H. J., McHugo, G. J., & Wolfe, R. (2008). A randomized controlled trial of cognitive-behavioral treatment for posttraumatic stress disorder in severe mental illness. *Journal of Consulting and Clinical Psychology, 76,* 259–271.

Munro, M. D. (2006). *Client and therapist variability with psychotherapy homework : A preliminary psychometric evaluation of two scales.* Unpublished Master's Thesis, Massey University, Albany, New Zealand.

Muran, J. C., Safran, J. D, Gorman, B. S., Wallner-Samstag, L., Eubanks-Carter, C., & Winston, A. (2009). The relationship of early alliance ruptures and their resolution to process and outcome in three time-limited psychotherapies for personality disorders. *Psychotherapy, Theory, Research, Practice, Training, 46,* 233-248.

Muran, J. C., Segal, Z. V., Samstag, L., & Crawford, C. (1994). Patient pretreatment interpersonal problems and therapeutic alliance in short-term cognitive therapy. *Journal of Consulting and Clinical Psychology, 62,* 185-190.

Murray, C. J. L., & Lopaz, A. D. (1996). *The global burden of disease: A comprehensive assessment of mortality and disability from diseases, injuries and risk factors in 1990 and projected to 2020.* Cambridge: Published by the Harvard School of Public Health on behalf of The World Health Organization and The World Bank. Distributed by Harvard University Press.

Needles, D. J., & Abramson, L. Y. (1980). Positive life events, attribution style, and hopefulness: Testing a model of recovery from depression. *Journal of Abnormal Psychology, 99,* 156-165.

Neimeyer, R. A. (1993). An appraisal of constructivist psychotherapies. *Journal of Consulting and Clinical Psychology, 61,* 221-234.

Newman, C. (1998). The therapeutic relationship and alliance in short-term cognitive therapy. In J. D. Safan & J. C. Muran (Eds.), *The therapeutic alliance in brief psychotherapy* (pp. 95-122). Washington, DC: American Psychological Association.

Newman, M., Stiles, W. B., Janeck, A., & Woody, S. R. (2006). Integration of therapeutic factors in anxiety disorders. In L. G. Castonguay & L. E. Beutler (Eds.). *Principles of therapeutic change that work* (pp.187-200). Oxford: Oxford University Press.

Nezu, A. M., & Nezu, C. M. (2005). Comments on "Evidence based behavioral medicine: What is it and How do we achieve it?": The interventionist does not always equal the intervention-the role of the therapist competence. *Annals of Behavioral Medicine, 29,* 80-81.

Nisbett, R.E., & Wilson, T.D. (1977). The halo effect: Evidence for the unconscious alteration of judgments. *Journal of Personality and Social Psychology, 35,* 450-456.

Norcross, J. C., (2002). Empirically supported therapy relationships. In J.C., Norcross (Ed.), Psychotherapy relationships that work: *Therapist contributions and responsiveness to patients* (pp. 3-16). New York: Oxford University Press.

Nott, L., Reeve, C., & Reeve, R. (1992). Scoring rubrics: An assessment option. *Science Scope, 15,* 44-45.

Oakley Browne, M. A., Wells, J.E., & Scott, K. M. (2006).*Te Rau Hinengaro: The New Zealand Mental Health Survey*. Wellington: Ministry of Health.

O'Farrell, M. K., Hill, C. E., & Patton, S. (1986). Comparison of two cases of counseling with the same counselor. *Journal of Counseling and Development, 65,* 141-145.

Ogrodniczuk, J. S., Piper, W. E., Joyce, A. S., & McCallum, M. (2000). Different perspectives of the therapeutic alliance and therapist technique in 2 forms of dynamically oriented psychotherapy. *Canadian Journal of Psychiatry, 45,* 452-458.

Orlinsky, D. E., & Howard, K. I. (1986a). Process and outcome in psychotherapy. In S. L. Garfield & A. E. Bergin & (Eds.), *Handbook of psychotherapies and behavior change* (3rd ed., pp. 311- 381). New York: John Wiley & Sons.

Orlinsky, D. E., & Howard, K. I. (1986b). The psychological interior of psychotherapy: Explorations with the therapy session reports. In L. S. Greenberg & W. M. Pinsof (Eds.), *The psychotherapeutic process: A research handbook*. New York: Guilford Press.

Orlinsky, D. E., & Howard, K. I. (1995). Unity and diversity among psychotherapies: A comparative perspective. In B. Bongar & L. E. Beutler (Eds.), *Comprehensive textbook of Psychotherapy: Theory and practice* (pp. 3-23). New York: Oxford University Press.

Pachankis, J. E., & Goldfried, M. R. (2007). On the next generation of process Research. *Clinical Psychology Review, 27,* 760-768.

Padesky, C. A. (1993). *Socratic questioning: Changing minds or guided discovery?* A keynote address delivered at the European Congress of Behavioural and Cognitive Therapies. London: September 24th 1993.

Padesky, C. A. (2004). Aaron T. Beck: Mind, man and mentor. In R. L. Leahy (Ed.), *Contemporary cognitive therapy: Theory, research and practice* (pp. 3-24). New York: Guilford.

Padesky, C. A. & Greenberger, D. (1995). *Clinician's guide to mind over mood.* New York: The Guilford Press.

Pampallona, S., Bollini, P., Tibaldi, G., Kupelnick, B., & Munizza, C. (2004). Combined Pharmacology and psychological treatment for depression: A systematic review. *Archives of General Psychiatry, 61,* 714-719.

Parsonson, B. S., & Baer, D. M. (1978). The analysis and presentation of graphic data. In T. R. Kratochwill (Ed.), *Single-subject research: Strategies for evaluating change* (pp. 101-165). New York: Academic Press.

Perepletchikova, F., Treat, T. A., & Kazdin, A. E. (2007). Treatment integrity in psychotherapy research: Analysis of the studies and examination of the associated factors. *Journal of Consulting and Clinical Psychology, 75,* 829-841.

Persons, J. B. (1989). Cognitive therapy in practice: a case formulation approach. New York: Norton.

Persons, J. B., & Bertagonolia, A. (1999). Inter-rater reliability of cognitive-behavioral case formulations of depression: A replication. *Cognitive Therapy and Research, 23,* 271-283.

Persons, J. B., & Burns, D. D. (1985). Mechanisms of action of cognitive therapy: The relative contributions of technical and interpersonal interventions. *Cognitive Therapy and Research, 9,* 539-551.

Peterson, C., Semmel, A., von Baeyer, C., Abramson, L. Y., Metalsky, G. I., & Seligman, M. E. P. (1982). The attributional Style Questionnaire. *Cognitive Therapy and Research, 6*, 287-299.

Petry, N. M., Tennen, H., & Affleck, G. (2000). Stal king the elusive client variable in psychotherapy research. In C. R. Snyder & R. E. Ingram (Eds.), *Handbook of psychological change: Psychotherapy processes and practices for the 21st century* (pp. 88- 108). New York: John Wiley &Sons.

Pilling, S., Bebbington, D., Knipers, E., Garety, P., Geddes, J., Orbach G., & Morgan, C. (2002). Psychological treatments in schizophrenia: II. Meta-analysis of family and cognitive-behavior therapy. *Psychological Medicine, 35,* 763-782.

Pinsoff, W. M. (1994) An integrative systems perspective on he therapeutic alliance: Theoretical clinical, and research limitations. In A. O. Horvath & L. S. Greenberg (Eds.), *The working alliance: Theory, research and practice* (pp. 173-195). New York: John Wiley & Sons.

Piper, W.E., Azim, H. F., Joyce, A. S., McCallum, M., Nixon, G. W. H., & Segal, P. S. (1991). Quality of object relations versus interpersonal functioning as predictors of of therapeutic alliance and psychotherapy outcome. *Journal of Nervous and Mental Disease, 179*, 432-438.

Plotnicov, K.H. (1990). *Early termination from counseling: The client's perspective.* Unpublished doctoral thesis, University of Pittsburgh, PA.

Portney, L.G., & Watkins, M. P. (1993). *Foundations of clinical research, Applications and practice.* Norwalk, Conneticut: Appleton & Lange.

Pos, A. E., Greenberg, L. S., Goldman, R. N., & Korman, L. M. (2003). Emotional processing during experiential treatment of depression. *Journal of Consulting and Clinical Psychology, 71,* 1007-1016.

Principe, J. M., Marci, C. D., Glick, D. M., & Ablon, J. S. (2006). The relationship among patient contemplation, early alliance, and continuation in psychotherapy. *Psychotherapy: Theory, Research, Practice, Training, 43,* 238-242.

Pulakos, E. D. (1986). The development of training programs to increase accuracy with different training tools. *Organizational Behavior and Human Decision Processes, 38,* 76-91.

Rachman, S. (1996). Trends in cognitive and behavioural therapies. In P.Salkovskis (Ed.), *Trends in cognitive and behavioural therapies* (pp. 1-24). Oxford, UK: John Wiley & Sons.

Rachman, S J., & Wilson, G. T. (1980). *The effects of psychological therapy.* Pergamon Press, Toronto.

Ramana, R., Paykel, E. S., Cooper, Z., Hayhurst, H., Saxty, M., & Surtees, P. G. (1995). Remission and relapse in major depression: a two-year prospective follow-up study. *Psychological Medicine, 25,* 1161-1170.

Raue, P. J., (1995) *Parameters of the therapeutic alliance in psychodynamic-interpersonal and cognitive-behavior therapy.* Unpublished doctoral thesis, State University of New York at Stony Brook. New York.

Raue, P. J., Castonguay, L. G., & Goldfried, M. R. (1993). The working alliance: A comparison of two therapies. *Psychotherapy Research, 3,* 197-207.

Raue, P. J., Castonguay, L. G., Newman, M. Gaus-Binkley, V., Shearer, D., & Goldfried M. R. (1991). *Guidelines for the Working Alliance Inventory–Observer Form (WAI-O).* Unpublished manuscript, State University of New York at Stony Brook. New York.

Raue, P. J., & Goldfried, M. R. (1994). The therapeutic alliance in cognitive-behavior therapy. In A. O. Horvath & L. S. Greenberg (Eds.), *The working alliance: Theory, research and practice* (pp. 131-152). New York: John Wiley & Sons.

Raue, P. J., Goldfried, M. R., & Barkham, M. (1997). The therapeutic alliance in psychodynamic-interpersonal and cognitive behavioral therapy. *Journal of Consulting and Clinical Psychology, 65,* 582-587.

Raza, S. M., & Carpenter, B. N. (1987). A model of hiring decisions in real employment interviews. *Journal of Applied Psychology, 72,* 596-603.

Reandeau, S. G., & Wampold, B. E. (1991). Relationship of power and involvement to working alliance: A multiple case sequential analysis of brief therapy. *Journal of Counseling Psychology, 38,* 107-114.

Rector, N. A., Zuroff, D. C., & Segal, Z. V. (1999). Cognitive change and the therapeutic alliance: The role of technical and nontechnical factors in cognitive therapy. *Psychotherapy, 36,* 320-328.

Reichelt, K. F., James, I. A., & Blackburn, I. (2003). Impact of training on rating competence in cognitive therapy. *Journal of Behavior Therapy and Experiential Psychiatry, 34,* 87-99.

Reid, R. B. (1970). Reliability assessment of observation data: A possible methodological problem. *Child Development, 41,* 1143-1150.

Reinecke, M. A. & Freeman, A. (2003) Cognitive therapy. In A. Gurman & S. Messer (Eds.) *Essential psychotherapies: Theory and Practice* (2nd ed, pp. 224-271). New York: Guilford Press.

Richmond, R. (1992). Discriminating variables among psychotherapy dropouts from a psychological training clinic. *Professional Psychology: Research and Practice, 24,* 190-195.

Rogers, C. R. (1951). *Client – centered therapy.* London: Constable and Company Ltd.

Rogers, C. R. (1957). The necessary and sufficient conditions of therapeutic personality change. *Journal of Consulting and Clinical Psychology, 21,* 95-103.

Romanczyk, R. G., Kent, R. N. Diament, C., & O'Leary, K. D. (1973). Measuring the reliability of observational data: A reactive process. *Journal of Applied Behavioral Analysis, 6,* 175-184.

Rosenman, S. J., Korten, A. E., & Levings, C. T. (1997). Computerised diagnosis in acute psychiatry: Validity of CIDI-Auto against routine clinical diagnosis. *Journal of Psychiatric Research, 31,* 581-592.

Rosner, B. (2006).*Fundamentals of biostatistics,* (6th ed). Belmont, CA: Duxbury Press.

Roth, A., & Fonagy, P. (1996). *What works for whom? A critical review of psychotherapy research.* London: Guilford Press.

Roth, A. & Fonagy, P. (2005). *What works for whom? A critical review of psychotherapy research* (2nd ed.). London: Guilford Press.

Rudner, L.M. (1992). Reducing errors due to the use of judges. Practical Assessment, Research & Evaluation, 3(3). Retrieved March 20, 2009 from http://PAREonline.net/getvn.asp?v=3&n=3.

Saal, F, R., Downey, R. G., & Lahey**,** M. A. (1980). Rating the ratings: assessing psychometric properties of rating data. *Psychological Bulletin, 88,* 413-428.

Saal, F. R., & Landy, F. J. (1977). The mixed standard rating scale: An evaluation. *Organizational Behavior and Human Performance, 18, 19-35.*

Sadock, B. J. & Sadock, V. A. (2003). *Kaplan & Sadock's synopsis of psychiatry: Behavioral sciences/clinical psychiatry* (9th ed.). New York: Lippincott Williams & Wilkins.

Safran, J. D., Crocker, P., McMain, S., & Murray, P. (1998). The therapeutic alliance rupture as a therapy event for empirical investigation. In J. D. Safran (Ed.), *Widening the scope of cognitive therapy: the relationship, emotion, and the process of change* (pp. 187-209). Northvale, New Jersey: Jason Aronson Inc.

Safran, J. D., & Muran, J. C. (2000b*). Negotiating the therapeutic alliance: A relational treatment guide.* New York: Guilford Press.

Safran, J. D., Muran, J. C., Samstag, L. W., & Stevens, C. (2002). Repairing alliance rupture. In J.C. Norcross (Ed.), *Psychotherapy relationships that work: Therapist contributions and responsiveness to patients (*pp.235-254). Oxford: Oxford University Press.

Safran, J. D., & Segal, Z. V. (1996). *Interpersonal process in cognitive therapy.* New York: Basic.

Safran, J. D., Segal, Z. V., Vallis, T. M., Shaw, B. F., & Samstag, L. W. (1993). Assessing patient suitability for cognitive therapy with an interpersonal focus. *Cognitive Therapy and Research 17,* 23-28.

Safran, J. D., & Wallner, L. K. (1991). The relative predictive validity of two therapeutic measures in cognitive therapy. *Psychological Assessment, 3,* 188-195.

Salkovskis, P.M. (Ed.). (1996). *Frontiers of cognitive therapy.* New York: Guilford Press.

Samstag, L.W. (2006). The working alliance in psychotherapy: An overview of the invited papers in the special section. *Psychotherapy: Theory, Practice, Training, 43*, 300-307.

Schneider, K. J. (2001). Multiple-case depth research: Bringing experience-near closer. In K. J. Schneider, J. F. T. Bugental, & J. F. Pierson (Eds.), *The handbook of humanistic psychology: leading edges in theory, research, and practice* (pp. 305-314). California: Sage Publications.

Schneider, K. J. (2003). Existential-humanistic psychotherapies. In A. S. Gurman & S. B. Messer (Eds.), *Essential psychotherapies: Theory and practice* (2nd ed, pp. 149-181). New York: The Guilford Press.

Schoevers, R. A., Beekman, A. T. F., Deeg, D. J. H., Jonker, C., & van Tilburg, W. (2003). Comorbidity and risk-patterns of depression, generalised anxiety disorder and mixed anxiety-depression in later life: results from the AMSTEL study. *International Journal of Geriatric Psychiatry, 18,* 994-1001.

Scott, J. (2001). Cognitive therapy for depression. *British Medical Bulletin, 57,* 101-113.

Scott, J., Garland, A., & Moorehead, S. (2001). A pilot study of cognitive therapy in bipolar disorder. *Psychological Medicine, 31*, 459-467.

Scott, J., Palmer, S., Paykel, E., Teasdale, J., & Hayhurst, H. (2003). Use of cognitive therapy for relapse prevention in chronic depression. *British Journal of Psychiatry, 182,* 21-227.

Sexton, H. (1996). Process, life events, and symptomatic change in brief eclectic Psychotherapy. *Journal of Consulting and Clinical Psychology, 64,* 1358-1365.

Sexton, H., Littauer, H., Sexton, A., & Tommeras, E. (2005). Building the alliance: Early therapeutic process and the client-therapist connection. *Psychotherapy Research, 15,* 103-116.

Shapiro, J. G. (1968). Relationships between expert and neophyte ratings of therapeutic conditions. *Journal of Consulting and Clinical Psychology, 32,* 87-89.

Shohamy, E., Gordon, C. M., & Kraemer, R. (1992). The effect of raters' background and training on the reliability of direct writing tests. *The Modern Language Journal, 76,* 27-33.

Shrout, P. E., & Fleiss, J. L. (1979). Intraclass correlations: Uses in assessing rater Reliability. *Psychological Bulletin, 86,* 420-428.

Sidman, M. (1960). *Tactics of scientific research; evaluating experimental data in psychology.* New York: Basic Books.

Siegart, R. J., Walkey, F. H., & Taylor, A. J.(1992) Anxiety, depression and cognitive self-statements: A factor analytic study. *New Zealand Journal of Psychology, 21,*

Sledge, W.H., Moras, K., Hartley, D., & Levine, M. (1990). Effect of time-limited therapy on client drop-out rates. *American Journal of Psychiatry, 147,* 1341-1347.

Smith, T. L., Barrett, M. S., Benjamin, L. S., & Barber, J. P. (2006). Relationship factors in treating personality disorders. In L. G. Castonguay & L. E. Beutler (Eds.), *Principles of therapeutic change that work* (pp. 219-238). Oxford: Oxford University Press.

Snyder, C., Michael, S., & Cheavens, J. (1999). Hope as a psychotherapeutic foundation of common factors, placebos and expectancies. In M. Hubble, D.Duncan & S. Miller (Eds.), *The heart and soul of change: What works in therapy.* Washington, D C: American Psychological Association.

Souery, D., Oswald, P., Massat, I., Bailer, U., Bollen, J., Demyttenaere, K., Kasper, S., Lecrubier, Y., Montgomery, S., Serretti, A., Zohar, J., Mendlewicz, J. (2007). Clinical factors associated with treatment resistance in major depressive disorder: Results from a European multicenter study. *Journal of Clinical Psychiatry, 68*, 1062-1070.

Staples, F. R., Sloane, R. B., Whipple, K., Cristol, A. H., & Yorkston, N. J. (1975). Differences between behavior therapists and psychotherapists. *Archives of General Psychiatry, 32,* 1517-1522.

Sterba, R. F. (1934). The fate of the ego in analytic therapy. *International Journal of Psychoanalysis, 115, 117*-224.

Stiles, W. B., Agnew-Davies, G. E., Barkham, Culverwell, A., Goldfried, M. R.,Halstead J., Hardy, G.E., Raue, P.J., Rees, A., & Shapiro, D.A. (2002). Convergent validity of the Agnew Relationship Measure and the Working Alliance Inventory. *Psychological Assessment 14*, 200-220.

Stiles, W. B., Agnew-Davies, R., Hardy, G. E., Barkham, M., & Shapiro, D. A. (1998). Relations of the alliance with psychotherapy outcome: Findings from the second Sheffield psychotherapy project. *Journal of Consulting and Clinical Psychology, 66,* 791-802.

Strauss, J. L., Hayes, A. M., Johnson, S. L., Newman, C. F., Brown, G. K., Barber, J. P., Laurenceau, J., & Beck, A.T. (2006). Early alliance, alliance ruptures, and symptom change in a nonrandomized trial of cognitive therapy for avoidant and obsessive-compulsive personality disorders. *Journal of Consulting and Clinical Psychology, 74,* 337-345.

Sweet, A. A. (1984). The therapeutic relationship in behavior therapy. *Clinical Psychology Review, 4, 253*-272.

Taft, C. T., Murphy, C. M., King, D. W., Musser, P. H., & DeDeyn, J. M. (2003). Process and treatment adherence factors in group cognitive-behavioral therapy for partner violent men. *Journal of Consulting & Clinical Psychology, 71,* 812-820.

Tang, T. Z., & DeRubeis, R. J. (1999). Sudden gains and critical sessions in cognitive- behavioral treatment for depression. *Journal of Consulting and Clinical Psychology, 6,* 894-904.

Tang, T. Z., DeRubeis, R. J., Beberman, R., & Pham, T. (2005). Cognitive changes, critical sessions and sudden gains in cognitive-behavioral treatment for depression. *Journal of Consulting and Clinical Psychology, 73,* 168-172.

Talley, P., Strupp, H., & Morey, L. (1990). Matchmaking in psychotherapy: Patient-therapist diamensions and their impact on outcome. *Journal of Consulting and Clinical Psychology, 58,* 182-188.

Thompson, B. (2003). Understanding reliability and coefficient alpha, really. In B. Thompson (Ed).), *Score reliability: Contemporary thinking on reliability issues.* Thousand Oaks, CA: Sage Productions.

Thorndike, E. L. (1920). A constant error in psychological ratings. *Journal of Applied Psychology, 4,* 25-29.

Tichenor V., & Hill, C. E. (1989). A comparison of six measures of working alliance. *Psychotherapy, 26*, 195-199.

Tinsley, H., & Weiss, D. (1975). Interrater reliability and agreement of subjective judgments. *Journal of Counseling Psychology, 22*, 358-376.

Tompkins, A. M. (2003). Effective homework. In R.L. Leahy (Ed.), *Roadblocks in cognitive – behavioural therapy: Transforming challenges into opportunities for change* (pp. 49-69). New York: Guilford Press.

Tracey, T. J. (1993). An interpersonal stage model of the therapeutic process. *Journal of Counseling Psychology, 40,* 396-409.

Tracey, T. J., & Kokotovic, A. M. (1989). Factor structure of the Working Alliance Inventory. *Psychological Assessment: A Journal of Consulting and Clinical Psychology, 1,* 207-210.

Treasure, J.L., Katzman, M., Schmidt, U., Troop, N., Todd, G., & de Silva, P. (1999). Engagement and outcome in the treatment of bulimia nervosa: First phase of a sequential design comparing motivation enhancement therapy and cognitive behavioural therapy. *Behaviour, Research and Therapy, 37,* 405-418.

Tryon, G. S. (1990). Session depth and smoothness in relation to the concept of engagement in counseling. *Journal of Counseling Psychology, 37,* 248-263.

Tryon, G. S., Blackwell, S. C., & Hammel, E. F. (2008). The magnitude of client and therapist alliance ratings. *Psychotherapy, Theory, Research Practice, Training, 45,* 546-551.

Tryon G. S., & Kane, A. S. (1995). Client involvement, working alliance, and type of therapy termination. *Psychotherapy Research, 5,* 189-198.

Tryon, G. S. & Winograd, G. (2002). Goal consensus and collaboration. In J.C. Norcross (Ed.), *Psychotherapy relationships that work: Therapist contributions and responsiveness to patients* (pp.109-125). Oxford: Oxford University Press.

Tsui, A. S. (1983). *Qualities of judgmental ratings by four rater sources.* Paper presented at the annual conference of the American Psychological Association, Anaheim, CA. (ERIC Document Reproduction Service no. ED 237913).

Uebersax, J. S. (2003). Statistical methods for rater agreement. Retrieved 25 February, 2008 from http://ourworld.compuserve.com/homepage/jsuebersax/agree.htm

Valentine, H. (2005). *How important is treatment integrity in randomised controlled trials dealing with cognitive behavioural therapy?* Unpublished Literature Review.

Vallis, T. M., Shaw, B. F., & Dobson, K. S. (1986). The Cognitive Therapy Scale: Psychometric properties. *Journal of Consulting and Clinical Psychology, 54,* 381-385.

Vance, R.J., Winne, P.S., & Wright, E.S. (1983). A longitudinal examination of rater and rate effects in performance ratings. *Personnel Psychology, 36,* 609-620.

Van Londen, L., Molenaar, R. P. G., Goekoop, J. G., Zwinderman, A. H., & Rooijmans, H. G. M. (1998). Three-to-5 year prospective follow-up of outcome in major depression. *Psychological Medicine, 28,* 731-735.

van Oppen, P., de Haan, E., van Balkom, A., Spinhoven, P., Hoogduin, K., & van Dyck, R. (1995). Cognitive therapy and exposure in vivo in the treatment of obsessive compulsive disorder. *Behavior Research and Therapy, 33*, 379-390.

Vincent, A.H. (2002). Measurement issues in determining interrater agreement (Doctoral Thesis, University of Oklahoma, 2002). *Thesis Abstract International,* 63(3-B), 1607. (UMI No. AAT 3045846).

Vittengl, J. R, Clark, L. A., Dunn, T. W., & Jarrett, R. B. (2007). Reducing relapse and reoccurrence in unipolar depression: A comparative meta-analysis of cognitive- behavioral therapy's effects. *Journal of Consulting and Clinical Psychology, 75,* 475-488.

von Raffler-Engel, W. (1980). *The perception of nonverbal behavior in the function of the age and the sex of the rater.* Paper presented at the annual meeting of the International Congress of psychology, Leipzig, Germany.

Waddington, L. (2002). The therapeutic relationship in cognitive therapy: A review. *Behavioural and Cognitive Psychotherapy, 30,* 179-191.

Walsh, R. A., & McElwain, B. (2002). Existential psychotherapies. In J. J. C. Cain & J. Seeman (Eds.), *Humanistic psychotherapies: Handbook of research and practice* (pp. 253-278). Washington, DC: American Psychological Association.

Walter, S. D., Eliasziw, M., & Donner, A. (1998). Sample size and optimal designs for reliability studies. *Statistics in Medicine, 17,* 101-110.

Waltz, J., Addis, M. E., Koerner, K., & Jacobson, N. S. (1993). Testing the integrity of a psychotherapy protocol: Assessment of adherence and competence. *Journal of Consulting and Clinical Psychology, 61, 620-630.,*

Wampold, B. E. (1997). Methodological problems in identifying efficacious psychotherapies. *Psychotherapy Research, 7,* 21-43.

Wampold, B. E. (2001). *The great psychotherapy debate: Models, Methods, and findings.* Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Wampold, B. E. (2007). Psychotherapy: The humanistic (and effective) treatment. *American Psychologist, 62,* 857-873.

Wampold, B. E., Minami, T., Baskin, T. W., & Callen Tierney, S. (2002). A meta (re)analysis of the effects of cognitive therapy versus "other therapies" for depression. *Journal of Affective Disorders, 68,* 159-165.

Warmke, D. L., & Billings, R. S. (1979). Comparison of training methods for improving the psychometric quality of experimental and administrative performance ratings. *Journal of Applied Psychology, 64,* 124-131.

Watson, J. C. (2001). Revisioning empathy: Theory, research and practice. In D. Cain & J. Seeman (Eds.). *Humanistic psychotherapies: Handbook of research and practice* (pp. 445-472). Washington DC: American psychological Association.

Watson, J. C. (2007). Reassessing Rogers' necessary and sufficient conditions of change. *Psychotherapy: Theory, Research, Practice, Training, 44,* 268-273.

Watson, J. C., & Geller, S. M. (2005). The relation between the relationship conditions, working alliance, and outcome in both process-experiential and cognitive-behavioral psychotherapy. *Psychotherapy Research, 15,* 25-33.

Watson, J. C., & Greenberg, L. S. (1994). The alliance in experiential therapy; Enacting the relationship conditions. In A. O. Horvath & L. S. Greenberg (Eds.), *The working alliance: Theory, research and practice* (pp.153-172). New York: John Wiley & Sons, Inc.

Watson, J. C., & Greenberg, L. S. (1998). The therapeutic alliance in short-term humanistic and experiential therapies. In J. D. Safran & J. C. Muran (Eds.), *The therapeutic alliance in brief therapies* (pp.123-145). Washington, DC: American Psychological Association.

Watson, J. C., Greenberg. L. S., & Lietaer, G. (1998). The experiential paradigm unfolding: Relationship and experiencing in theory. In L. S Greenberg, J. C. Watson & G. Lietaer (Eds), *Handbook of experiential psychotherapy.* (pp. 3-27). New York: The Guilford Press.

Watson, J. C., & McMullen, E. J. (2005). An examination of therapist and client behavior in high- and low-alliance sessions in cognitive-behavioral therapy and process experiential therapy. *Psychotherapy: Theory, Research, Practice, Training, 42,* 297-310.

Wedge, P. (2005). The development of the Homework Adherence and Competence Scale (HAACS): A measure for assessing therapist adherence and competence in administrating homework assignments within cognitive behaviour therapy. Unpublished Master's Thesis, Massey University, Albany, New Zealand.

Wells, K. B., Burnam, M. A., Rogers, W., Hays, R., & Camp, P. (1992). The course of depression in adult outpatients. Results from the Medical Outcomes Study. *Archives of General Psychiatry, 49,* 788-794.

Westen, D., & Morrison, K. (2001). A multi-dimensional meta-analysis of treatments for depression, panic and generalized anxiety disorder: an empirical examination of the status of empirically supported therapies. *Journal of Consulting and Clinical Psychology, 69,* 319-328.

Whisman, M. A. (1993). Mediators and moderators of change in cognitive therapy of depression. *Psychological Bulletin, 114,* 248-265.

Whisman, M. A. (1999). The importance of cognitive theory of change in cognitive therapy of depression. *Psychology: Science and Practice, 6,* 300-304.

Wierzbki, M., & Pekarik, G. (1993). A meta-analysis of psychotherapy dropout. *Professional Psychology: Research and Practice, 24,* 190-195.

Wilson, G. T., Fairburn, C. C., Agras, W. S., Walsh B. T., & Kraemer, H. (2002). Cognitive-behavioral therapy for bulimia nervosa: Time course and mechanisms for change. *Journal of Consulting and Clinical Psychology, 70,* 267-274.

Wilson, G. T., Loeb, K. L., Walsh, B. T., Labouvie, E., Petkova, E., Liu, X., & Waternaux, C. (1999). Psychological versus pharmacological treatments of bulimia nervosa: predictors and processes of change. *Journal of Consulting and Clinical Psychology, 67,* 451-459.

Wittchen, H. U., (1994). Reliability and validity studies of the WHO-Composite International Diagnostic Interview (CIDI): A critical review. *Journal of Psychiatric Research, 28,* 57-84.

Wolfe, B. E., & Goldfried, M. R. (1988). Research on psychotherapy integration: Recommendations and conclusions from NIMH workshop. *Journal of Consulting and Clinical Psychology, 56,* 448-451.

Woody, S. R., & Adessky, R. S. (2002). Therapeutic alliance, group cohesion, and homework compliance during cognitive-behavioral group treatment of social phobia. *Behavior Therapy*, 33, 5-27.

World Health Organisation. (1997). *Composite International Diagnostic Interview (CIDI): Version 2.1.* Geneva: World Health Organisation.

Wright, J. H., & Davis, D. (1994). The therapeutic relationship in cognitive-behavioral therapy: Patient perceptions and therapist responses. *Cognitive and Behavioral Practice, 1,* 25-45.

Young, J. E., & Beck, A. T. (1980) *Cognitive Therapy Scale Rating Manual.* (Available from the Beck Institute for Cognitive Therapy and Research, GSB Building, One Belmont Avenue, Suite 700, Bala Cynwyd, PA 19004-1610).

Young, J. E., Rygh, J. L., Weinberger, A. D., & Beck, A. T. (2008). Cognitive therapy of depression. In D. Barlow (Ed.), *Clinical handbook of psychological disorders: A step-by-step treatment manual* (4th ed. pp. 206-244). New York: Guilford Press.

Zedeck, S., & Cascio, W. F. (1982). Performance appraisal decisions as a function of rater training and purpose of the appraisal. *Journal of Applied Psychology, 67,* 752-758.

Zetzel, E. (1956). Current concepts of transference. *The International Journal of Psycho-Analysis, 15,* 117-123.

Zimmermann, G., Favrod, J., Trieu, V. H., & Pomini, V. (2005). The effect of Cognitive behavioral treatment on the positive symptoms of schizophrenia spectrum disorders: A meta-analysis. *Schizophrenia Research, 77,* 1-9.

Zuroff, D. C., & Blatt, S. J. (2006). The therapeutic relationship in the brief treatment of depression: Contributions to clinical improvement and enhanced capacities. *Journal of Consulting and Clinical Psychology, 74,* 130-140.

Zuroff, D. C., Blatt, S. J., Sotsky, S. M., Krupnick, J. L., Martin, D. J., Sanislow, C. A., & Simmens, S. (2000). Relation of therapeutic alliance and perfectionism to outcome in brief outpatient treatment for depression. *Journal of Consulting and Clinical Psychology, 68,* 114-124.

APPENDICES

Appendix A

Participant Information and Consent Forms

Summary Information Sheet

A Depression Study

You are invited to take part in a research study involving a brief psychological treatment called Cognitive Behaviour Therapy (CBT). The study will involve 70 individuals between the age of 18 and 65 years, recruited within the Auckland area. (Like yourself,)These individuals will currently be experiencing a major depressive episode for the first time. Before you consent to be part of this study, please read the following. Ask as many questions as you need to be sure that you understand what taking part will involve. The decision to take part is entirely your choice, however due to time constraints it would be greatly appreciated if you could respond to this invitation within 1 week of receiving full information about the study.

After providing written consent, you will receive a 20 session protocol of CBT for depression over a 16 week period. Treatment will be individualised based on your specific needs and goals. Consistent with prior research on CBT for depression, the first 8 sessions will be offered twice weekly. Follow-up sessions will occur at 2 months and 6 months after treatment has ended. Participants will be asked to complete some assessment questionnaires to determine treatment gains, and also asked to provide informal feedback on the CBT they received. Therapy will be provided by clinical psychology doctoral/masters students trained in delivering this protocol.

How will the study benefit you? It is expected that new information, which may benefit you or others, will be obtained by this study. Furthermore, it is very likely that the comprehensive psychological assessment and therapy offered as part of this study will improve your condition, although this cannot be guaranteed. These services will be provided free of charge. Due to funding limitations, you will be responsible for your own travel costs to and from the Centre for Psychology in Albany, however parking will be provided free of charge.

Who is unable to take part? Participants will need to be proficient in reading, writing, and conversing in English. They must be free from taking drugs which act on the central nervous system. They must not meet diagnostic criteria for substance abuse, psychosis, or borderline personality disorder. Lastly, they must be able to be managed safely with outpatient psychotherapy.

If you do agree to take part you are free to withdraw from the study at any time without having to give a reason. This will in no way affect your continuing health care. Participation in this study will be stopped should any harmful effects appear or if an appropriate medical professional feels it is not in your best interest to continue. You may be taken out of the study if you need treatment that is not allowed during this study, or if the study is cancelled. You will be asked to check with your study therapist before taking any other treatment; this includes anything from the supermarket, pharmacy or health shop.
Will my information remain confidential? Participating in this study will involve having your therapy sessions videotaped (and transferred to DVD discs) in order for the researchers to monitor the therapy protocol. All information collected about you during the

study, including the recorded sessions, will be kept strictly confidential and only accessed by those researchers and clinical supervisors directly involved in the study. The only time in which confidentiality is breached is in the event that you express an intention to harm either yourself or somebody else, in which case a crisis team would become involved. No material which could personally identify you will be used in any reports on this study. All assessment information and clinical notes will be kept in individual files stored in a locked clinical records room, with files coded with anonymous identification numbers. Files will be stored in a separate location from both the identifying information and the DVD archive.

The information collected will be used for the research project and may be submitted for publication in an academic journal. All participants will be offered a summary of the findings at the conclusion of the study. This will include details of any publication arrangements that have been made. Please note that there is likely to be a delay between data collection and publication.

If at any time you have questions or concerns about this study, please contact:
Dr. Nik Kazantzis   Phone: (09) 414 0800 extension ……
                              Email: n.kazantzis@massey.ac.nz

If you have any queries or concerns regarding your rights as a participant in this study you may wish to contact a Health and Disability Advocate, telephone 0800 555 050 Northland to Franklin.

This study has received ethical approval from the Northern X Regional Ethics Committee. (NTX/06/07/085 - The Relationship Between Therapist Competence in Using Homework Assignments, Patient Homework Adherence, and Treatment Outcome in Cognitive Behaviour Therapy for Depression).

The Centre for Psychology Depression Study

PARTICIPANT CONSENT FORM

This consent form will be held for a period of five (5) years

I have read and I understand the Information Sheet dated 17 July, 2006, for volunteers taking part in the study designed to investigate the process of Cognitive Behaviour Therapy for depression, and have had the details of the study explained to me. I have had the opportunity to use whanau support or a friend to help me ask questions and understand the study. My questions have been answered to my satisfaction, and I understand that I may ask further questions at any time. I have been given contact details to use in case I have future questions about the study.

I understand that taking part in this study is voluntary (my choice) and that I may withdraw from the study at any time.

I understand that my participation in this study is confidential and that no material that could identify me will be used in any reports on this study.

I agree to my sessions in this study being video taped.

I understand that I will not receive any compensation for travel costs or for the time I spend as a participant in this study.

I have had adequate time to consider whether or not to take part in this study. I agree to participate in this study under the conditions set out in the Information Sheet.

Signature: .... **Date:** ....

Full Name - printed ....

APPENDIX B

WAI-SR-O

WAI-SR-O

Working Alliance Inventory – Short Form Revised – Observer Version

Name:

Session Number:

Date Session Rated:

Goal Items

1. The client and therapist collaborated on setting the goals for the session.

①	②	③	④	⑤
Seldom	**Sometimes**	**Fairly Often**	**Very Often**	**Always**

2. There is agreement on what is important for the client to work on.

①	②	③	④	⑤
Seldom	**Sometimes**	**Fairly Often**	**Very Often**	**Always**

3. The client and therapist are working on mutually agreed upon goals.

①	②	③	④	⑤
Seldom	**Sometimes**	**Fairly Often**	**Very Often**	**Always**

4. The client and therapist have established a good understanding of the changes that would be good for the client.

①	②	③	④	⑤
Seldom	**Sometimes**	**Fairly Often**	**Very Often**	**Always**

Task Items

5. There is agreement about the usefulness of the current activity in therapy (i.e., the client is seeing new ways to look at his/her problem).

①	②	③	④	⑤
Seldom	**Sometimes**	**Fairly Often**	**Very Often**	**Always**

6. There is agreement that what the client and therapist are doing in therapy will help the client to accomplish the changes he/she wants.

①	②	③	④	⑤
Seldom	**Sometimes**	**Fairly Often**	**Very Often**	**Always**

7. As a result of these sessions there is clarity about how the client might be able to change.

①	②	③	④	⑤
Seldom	**Sometimes**	**Fairly Often**	**Very Often**	**Always**

8. The client believes that the way they are working with his/her problem is correct.

①	②	③	④	⑤
Seldom	**Sometimes**	**Fairly Often**	**Very Often**	**Always**

Bond Items

9. There is a mutual liking between the client and therapist.

①	②	③	④	⑤
Seldom	**Sometimes**	**Fairly Often**	**Very Often**	**Always**

10. The client and the therapist respect each other.

①	②	③	④	⑤
Seldom	**Sometimes**	**Fairly Often**	**Very Often**	**Always**

11. The client feels the therapist respects and cares about the client, even when the client does things that the therapist does not approve of.

①	②	③	④	⑤
Seldom	**Sometimes**	**Fairly Often**	**Very Often**	**Always**

12. The client feels that the therapist appreciates him/her as a person.

①	②	③	④	⑤
Seldom	**Sometimes**	**Fairly Often**	**Very Often**	**Always**

Appendix C

Rating Guidelines for the WAI-SR-O

WAI-SR-O

Working Alliance Inventory
Short Form - Revised

Rating Guidelines for Supervisors & Independent Observers

Carol Osborne

Nikolaos Kazantzis, Ph.D.

First Draft – 13th February, 2008

Client Id Number:
Rater:
Session Number:
Session Date:
Date session rated:

WAI-SR-O

Instructions:

The Working Alliance Inventory-Short Form Revised (WAI-SR, Hatcher & Gillaspy, 2006) is intended as a clinical and research tool to be used by clinical supervisors and independent raters of audio and video recorded therapy sessions. The purpose of this measure is to rate the working alliance, based on three dimensions, (a) bond, (b) goals, and (c) tasks. The following is a general guide to rating these dimensions for a single therapy session, and are based on previous guidelines used with this measure (cf. Raue et al., 1991).

(a) **Goals**

Goals are defined as the objectives of the therapist and client that specifically target change. Goals can include reduction in symptoms, improvement in interpersonal skills or relationships, awareness of intrapersonal conflicts, and development of new ways of thinking or behaving. Agreement is assessed according to the extent to which both client and therapist see the goals as important, clear and capable of being accomplished. The understanding of the benefits resulting from the change and degree of confidence in change can also be assessed. It should also be noted that many of these examples of goals can actually be thought of as sub-goals or secondary goals. For example, increased activity may be desired in order to ultimately alleviate depression, whereas change in behavior toward ones partner may be a step toward improving the relationship.

(b) **Tasks**

Tasks are defined as therapeutic techniques or processes that help clients to increase awareness of their own thoughts, feelings, values and needs. Each task has a particular therapeutic goal. Examples of tasks are: support/reassurance (communicating an understanding of the clients internal state and a willingness to explore the client's thoughts and feelings), reflection (focusing on the deeper meaning of the clients speech in order to clarify feelings or facilitate their experience), reformulation (paraphrasing the clients speech in order to clarify experience, summarizing), cognitive restructuring (challenging illogical and maladaptive cognitions), any type of reality testing (pointing out evasions, resistance, self-deception of incongruities between speech and behavior), therapist suggestion /advice, information giving, homework assignments and role playing. Ratings on tasks are made according to how responsive the client is to task, and how responsive the therapist is to the client's need in suggesting the task. Therefore, the tasks should be seen as important, appropriate, and clear by therapist and client for high ratings. Of course, a specific technique, such as a particular intervention or homework assignment, may be rejected by the client without much effect on therapy overall. Agreement in these cases would depend on the degree of rejection (i.e., presence of thoughtful consideration of the technique) as well as their frequency.

(c) **Bond**

The therapeutic <u>bond</u> is defined as the mutual liking, respect, and trust between the client and the therapist. It is also characterized by therapist genuineness, openness, warmth, and understanding. Client confidence in the therapist, client comfort, and reciprocal respect are also included. The bond is assessed through the client's tone of voice, amount and quality (i.e., degree of comfort) of client talk concerning intimate issues. Bond is also assessed through the therapist degree of comfort, non-defensiveness, accurate empathy (especially empathy validated by the client), and the mutual value placed on each others contributions.

Please note that collaboration (i.e., both the therapist and the client participate in decision making, formulating ideas) in therapy requires agreement, whereas agreement (either the therapist or client suggests an idea/task/goal and the other agrees) is possible without collaboration.

Reference:
Raue, P. J., Castonguay, L.G., Newman, M., Gaus-Brinkley, V., Shearer, D., & Goldfried, M. R. (1991). Guidelines for the Working Alliance Inventory – Observer Form (WAI-O). Unpublished Manuscript, State University of New York at Stony Brook.

WAI-SR-O

Rating Guidelines;

- Ratings are based on frequency of occurrence within the session. However, intensity should also be taken into account. For example, one or two strong negative interactions between the therapist and client may characterize a session with a very negative alliance and should be rated as such. However, consider a higher rating if the rupture/strain is investigated and attempts are made to repair it. The intensity and frequency of rating the bond items vary in different ways to the goal and task items. Furthermore, the rating of task and goal items tends to place greater emphasis on frequency than bond items.

- All items refer to therapist and the client (dyad), but emphasis is placed on the client's experience. The most important criterion is the client's agreement with the therapist, belief in the therapists care and competence, and acceptance and participation in strategies used (e.g., willingly discloses, evaluates cognitions, or tries new behaviors).

- Rating the therapeutic alliance involves both subjective and objective evaluation criteria. The rater may have a personal reaction to the client and therapist that can color the rating of every item. On the other hand, the rater may rely solely on the observable in-session behavior. Ratings should be based on subjective reaction and observable behavior, that is, clinical judgment. For example, if the client appears positive to the therapist's concern, but **you** perceive the therapist as mechanical and emotionally distant from the client (possibly due to schema activation for the situation), then avoid interpreting the client's beliefs and rate the observable therapist behavior (i.e., this instance, a mechanical and emotionally distant style). The rationale is that the client's beliefs cannot always be inferred, and what an external observer perceives might, in fact, be affecting the client in some not immediate-observable way.

- However, ratings also need to be based on what is going on between the therapist and the client (do they seem to be working well together, have a good rapport, reciprocal relationship, the client is working and getting value out of the sessions). Care needs to taken that ratings are not based on the therapists style of relating.

- Consider all relevant information from a session. Try not to be distracted by elements that are particularly salient. For example, the client may disagree on some task or goal, yet come to an agreement later in the course of the session. Thus, take into account the information from the entire session.

- Do not consider the pathology or other characteristics of the client when making ratings. For example, a client may be non-disclosing and non-emotional by nature, but may (or may not) manage to open up with therapist help. Despite this improvement consider the case in relation to other cases and rate accordingly. Raters are often asked whether they should rate a client's level of alliance according to this client's range of capabilities or according to good alliance prototypes. The answer is that they should refer to realistic prototypes of good and poor alliances. Therefore ratings should not be idiosyncratic because the latter would prevent subject comparison.

- Do not rate the technical skill or competence of the therapist exhibited in the session. Although they may be related to competence, these ratings concern the therapeutic alliance. This rating scale places greater emphasis on timing, warmth, style, respect of client needs, than application of techniques or interventions.

- Ratings on this form assume the relationship quality is present. Subtract ratings when a therapist alliance behavior would have been helpful, when a therapist alliance behavior was clearly unhelpful, or when a therapist alliance behavior was absent. Simultaneously, the intensity or strength of the affective bond and the ability of the therapist and the client to work together in therapy needs to be taken into consideration. Thus, start with point '5' on each rating scale, and move down the scale as more therapeutic alliance elements are missing. For example, when rating "understanding" between the client and the therapist, the rater should move one point lower on the rating scale for any indication of misunderstanding, such as the therapist incorrectly reformulating the client's speech. Similarly, the rater should provide a lower rating if insufficient understanding was communicated through mechanical repetition of the client wording. A high score means a more intense and better alliance component.

1. Goal Item

①	②	③	④	⑤
Seldom	**Sometimes**	**Fairly Often**	**Very Often**	**Always**

The client and therapist <u>COLLABORATED ON SETTING THE GOALS</u> for the session

Collaboration on setting goals means both therapist and client work as a team to set the content of the in-session discussion or agenda. Collaboration is assessed by the responsiveness of the therapist and client to each other's initiations. Communication between the therapist and client is balanced. Points should be deducted if one participant changes the topic and the other continues to bring it up, even if the latter initially responds to the redirection. Where clients persistently bring up topics not agreed upon, consideration needs to be given to the therapist (a) politely interrupting and trying to bring client back on track, (b) discussing the client's reason for not wanting to keep to the agenda, or (c) re-negotiating session agenda. This item focuses on the actual collaboration process of setting and resetting goals.

Rating the intensity of the alliance maybe determined by whether the therapist and client are willingly and enthusiastically engaged in the therapeutic process of collaboratively setting goals for the session. There should be a sense that the client and the therapist think and feel that they are working toward mutually accepted goals. Ruptures can occur in the relationship when there is disagreement about therapy goals. For example, there may be frequent instances of the client and therapist collaboratively setting goals for the session, but if there appears to be little willingness or enthusiasm to take part in this process, and the reason for this lack of engagement is not discussed, a lower score may be considered.

5. There are frequent instances of the client and therapist collaboratively setting the goals for the session. There is a **VERY HIGH** level of engagement in the process of collaboratively setting goals for the session.

4. There are very often instances of the client and therapist collaboratively setting the goals for the session. However, there is **ONE** instance when collaboration did not take place. There is **a HIGH** level of engagement in the process of collaboratively setting goals for the session.

3. There are fairly often instances of the client and therapist collaboratively setting the goals for the session. However, there are **TWO** instances when collaboration did not take place. There is a **MODERATE** level of engagement in the process of collaboratively setting goals for the session.

2. There are sometimes instances of the client and therapist collaboratively set the goals for the session. However, there **THREE** instances when collaboration did not take place. There is a **LOW** level of engagement in the process of collaboratively setting goals for the session.

1. There are seldom instances of the client and therapist collaboratively set the goals for the session. There are **FOUR OR MORE** instances when collaboration did not take place. There is **VERY LOW, LITTLE, OR NO** engagement in the process of collaboratively setting goals for the session.

2. Goal Item

①	②	③	④	⑤
Seldom	**Sometimes**	**Fairly Often**	**Very Often**	**Always**

There is <u>AGREEMENT ON WHAT IS IMPORTANT</u> for the client to work on

Goals may refer to topics and concerns (in-session and long term goals) raised by the therapist and the client. This item relates to in session, between sessions and previous sessions. "Important" means therapeutically significant for the client. Agreement on what is important could include determining which problems are the most important to work on during the session/therapy, and the prioritizing of long term therapy goals and in-session agenda, Therapist monitors flow of discussion. If client and therapist inadvertently drift from the critical agenda item, consideration needs to be given to the therapist politely interrupting these peripheral discussions and returning both therapist and client to the agreed agenda item. Item two focuses more on agreement of what is important or content of the long term goals/in-session agenda/client problems.

Rating the intensity of the alliance maybe determined by whether the therapist and client are willingly and enthusiastically engaged in the therapeutic process of agreeing what is important for the client to work on. There should be a sense that the client and the therapist think and feel that they are working toward mutually accepted goals. Ruptures can occur in the relationship if there is disagreement about therapy goals. For example, there may be frequent instances of the client and therapist agreeing on what is important for the client to be working on, but if there appears to be little willingness or enthusiasm during this process, and the reason for this lack of engagement is not discussed, a lower score may be considered.

5. There are frequent instances of agreement on what is important for the client to work on. There is a **VERY HIGH** level of engagement in the process of agreeing on what is important for the client to be working on.

4. There are very often instances of agreement on what is important for the client to work on. However, there is **ONE** instance where more clarity could have been provided. There is a **HIGH** level of engagement in the process of agreeing on what is important for the client to be working on.

3. There are fairly often instances of agreement on what is important for the client to work on. However, there are **TWO** instances where more clarity could have been provided. There is a **MODERATE** level of engagement in the process of agreeing on what is important for the client to be working on.

2. There are sometimes instances of agreement on what is important for the client to work on. However, there are **THREE** instances where more clarity could have been provided. There is a **LOW** level of engagement in the process of agreeing on what is important for the client to be working on.

1. There are seldom instances of agreement on what is important for the client to work on. There are **FOUR OR MORE** instances where more clarity could have been provided. There is **VERY LOW, LITTLE, OR NO** engagement in the process of agreeing on what is important for the client to be working on.

191

3. Goal Item

①	②	③	④	⑤
Seldom	**Sometimes**	**Fairly Often**	**Very Often**	**Always**

The client and therapist are <u>WORKING ON MUTUALLY AGREED UPON GOALS</u>

This item refers more to in-session goals or agenda items. For example, the client and therapist work on agreed prioritized agenda items, therapist and client participate in 'in-session' activity which is linked to agenda/therapy goals, summaries are provided, and/or feedback and bridge provides transition between topics. Points should be deducted if the agenda is vague or incomplete, the therapist and client shift focus and the shift is accompanied by a redirection without resetting agenda or goals, the therapist is too directive, other issues are raised which are not on the agenda and the client and therapist have a discussion about that topic (e.g., a movie they had both seen). This item focuses on the therapist and the client actually working together on mutually agreed upon agenda items.

Rating the intensity of the alliance in this item maybe determined by whether the client and therapist are willingly and enthusiastically engaged in the therapeutic process of working on mutually agreed upon agenda goals. There should be a sense that the client and the therapist think and feel that they are working toward mutually accepted agenda goals. Ruptures can occur in the relationship if there is disagreement about agenda goals. For example, there may be frequent instances of the client and therapist working on mutually agreed upon agenda items, but if there appears to be little willingness or enthusiasm to take part in this process, and the reason for this lack of engagement is not discussed, a lower score may be considered.

5. There are frequently instances of the client and therapist working on mutually agreed upon goals. There is a **VERY HIGH** level of engagement in the process of the client and therapist working on mutually agreed upon goals.

4. There are very often instances of the client and therapist working on mutually agreed upon goals. However, there is **ONE** instance where there was no agreement. There is a **HIGH** level of engagement in the process of the client and therapist working on mutually agreed upon goals.

3. There are fairly often instances of the client and therapist working on mutually agreed upon goals. However, there are **TWO** instances where there was no agreement. There is a **MODERATE** level of engagement in the process of the client and therapist working on mutually agreed upon goals.

2. There are sometimes instances of the client and therapist working on mutually agreed upon goals. However, there are **THREE** instances where there was no agreement. There is a **LOW** level of engagement in the process of the client and therapist working on mutually agreed upon goals.

1. There are seldom instances of the client and therapist working on mutually agreed upon goals. There are **FOUR OR MORE** instances where there was no agreement. There is **VERY LOW, LITTLE, OR NO** engagement in the process of the client and therapist working on mutually agreed upon goals.

4. Goal Item

①	②	③	④	⑤
Seldom	**Sometimes**	**Fairly Often**	**Very Often**	**Always**

The client and therapist have established a <u>GOOD UNDERSTANDING OF THE CHANGES THAT WOULD BE GOOD FOR THE CLIENT</u>

Change implies that there is some degree of a focus on the future. The focus here should be on the client's goals for therapy and **what they might actually do** (the specific content) to reach these goals. For example, the client and therapist collaboratively agree on therapy goals, therapy goals are prioritized, a rationale is provided for interventions/homework assignments and this rationale is linked back to the client's therapy goals).

Rating the intensity of the alliance in this item maybe determined by therapist and client level of engagement when establishing a good understanding of the changes that would be good for the client. There should be a sense that the client and the therapist think and feel that they are working toward mutually accepted goals. Ruptures can occur in the relationship if there is disagreement about therapy goals. For example, there may be frequent instances where the client and therapist have established a good understanding of the changes that are good for the client, but the level of engagement in this process is low and the reason for this lack of engagement is not discussed, a lower score may be considered.

5. There are frequently instances of the client and therapist establishing a good understanding of the changes that would be good for the client. There is a **VERY HIGH** level of engagement between the therapist and client when establishing these changes.

4. There are very often instances of the client and therapist establishing a good understanding of the changes that would be good for the client. However, there is **ONE** instance where understanding was not established. There is a **HIGH** level of engagement between the therapist and client when establishing these changes.

3. There are fairly often instances of the client and therapist establishing a good understanding of the changes that would be good for the client. However, there are **TWO** instances where understanding was not established. There is a **MODERATE** level of engagement between the therapist and client when establishing these changes.

2. There are sometimes instances of the client and therapist establishing a good understanding of the changes that would be good for the client. However, there are **THREE** instances where understanding was not established. There is a **LOW** level of engagement between the therapist and client when establishing these changes.

1. There are seldom instances of the client and therapist establishing a good understanding of the changes that would be good for the client. There are **FOUR OR MORE** Instances where understanding was not established. There is a **VERY LOW, LITTLE, OR NO** engagement between the therapist and client when establishing these changes.

5. Task Item

①	②	③	④	⑤
Seldom	**Sometimes**	**Fairly Often**	**Very Often**	**Always**

There is agreement about the <u>USEFULNESS OF THE CURRENT ACTIVITY</u> in therapy (i.e., the client is seeing new ways to look at his/her problem).

This item refers to what is actually being worked on inside the current session. Disregard the parenthetical remark (which seems to place restrictions on possible strategies). Agreement about the usefulness of the current activity in therapy could include: agreement between client and therapist on about the rationale, importance, or usefulness of an intervention or homework assignment, and/or generalization of learning. The client understands the therapist's explanations. **Please note that collaboration in therapy requires agreement, whereas agreement is possible without collaboration.**

Intensity of the alliance in this item maybe determined by whether the therapist and client are willingly and enthusiastically engaged in task of agreeing about the usefulness of the current activity in therapy. The client is responsive to the task and the therapist is responsive to the clients needs for suggesting the task. Ruptures may occur in the relationship if there is disagreement about the task. For example, there may be frequent agreement about the usefulness of the current activity, but if there appears to be little willingness or enthusiasm to take part in the task and the reason for this lack of engagement is not discussed, a lower score may be considered.

5. There is frequent agreement about the usefulness of the current activity in therapy. There is a **VERY HIGH** level of engagement between the therapist and client when they discuss the usefulness of the current activity in therapy.

4. There are very often instances of agreement about the usefulness of the current activity in therapy. However, there is **ONE** instance where more clarity could be provided. There is a **HIGH** level of engagement between the therapist and client when they discuss the usefulness of the current activity in therapy.

3. There are fairly often instances of agreement about the usefulness of the current activity in therapy. However, there are **TWO** instances where more clarity could be provided. There is a **MODERATE** level of engagement between the therapist and client when they discuss the usefulness of the current activity in therapy.

2. There are sometimes instances of agreement about the usefulness of the current activity in therapy. However, there are **THREE** instances where more clarity could be provided. There is a **LOW** level of engagement between the therapist and client when discuss the usefulness of the current activity in therapy.

1. There are seldom instances of agreement about the usefulness of the current activity in therapy. There are **FOUR OR MORE** instances where more clarity could have been provided. There is a **VERY LOW, LITTLE OR NO** engagement between the therapist and client when they discuss usefulness of the current activity in therapy.

6. Task Item

①	②	③	④	⑤
Seldom	**Sometimes**	**Fairly Often**	**Very Often**	**Always**

There is <u>AGREEMENT</u> that what the client and therapist are doing in therapy will help the client to accomplish the changes he/she wants.

This item refers to in session, between session and previous session therapy. Examples for this item could include: agreement about chosen interventions and homework assignments, interventions and homework assignments are related to client goals, it is clear that the homework assignment is important to the client and that they are ready and confident to complete it, any barriers to complete an intervention or homework assignment are discussed.

Intensity of the alliance maybe determined by whether the therapist and client willingly and enthusiastically discussing what will help the client to accomplish these changes. The client is responsive to the task and the therapist is responsive to the clients needs for suggesting the task. For example, there may be frequent instances of agreement that what the client and therapist are doing in therapy will help the client to accomplish changes but if there appears to be little willingness or enthusiasm to take part in this task and the reason for this lack of engagement is not discussed, a lower score may be considered.

5. There are frequent instances of agreement that what the client and therapist are doing in therapy will help the client to accomplish the changes he/she wants. There is a **VERY HIGH** level of engagement between the therapist and client when they discuss these changes.

4. There are very often instances of agreement that what the client and therapist are doing in therapy will help the client accomplish the changes he/she wants. However, there is **ONE** instance where either the therapist or client expresses doubt or dissatisfaction and no other options are explored. There is a **HIGH** level of engagement between the therapist and client when they discuss these changes.

3. There are fairly often instances of agreement that what the client and therapist are doing in therapy will help the client accomplish the changes he/she wants. However, there are **TWO** instances where either the therapist or client expresses doubt or dissatisfaction and no other options are explored. There is a **MODERATE** level of engagement between the therapist and client when they discuss these changes

2. There are sometimes instances of agreement that what the client and therapist are doing in therapy will help the client accomplish the changes he/she wants. However, there are **THREE** instances where either the therapist or client expresses doubt or dissatisfaction and no other options are explored. There is a **LOW** level of engagement between the therapist and client when they discuss these changes.

1. There are seldom instances of agreement that what the client and therapist are doing in therapy will help the client accomplish the changes he/she want. There are **FOUR OR MORE** instances where either the therapist or client expresses doubt or dissatisfaction and no other options are explored. There is a **VERY LOW, LITTLE OR NO** engagement between the therapist and client when they discuss these changes.

7. Task Item

①	②	③	④	⑤
Seldom	**Sometimes**	**Fairly Often**	**Very Often**	**Always**

As a result of these sessions there is clarity about how the client might be able to CHANGE.

Change implies that there is some degree of a focus on the future. The focus here should be on **how the client might change** (the process of change/ways of dealing with problems) **not on what** the client might actually do (the specific content). This item is not about the content of the homework or intervention, but **how the client can do the task**. Discussion or assigning of homework this does not automatically rate a high rating. It is not enough that a thought record is mentioned by itself for homework, rather the homework task needs to have a rationale as to why it will help the client in the future and be specific (e.g., where the task will be completed, how often, when). What should be communicated is a sense of where the client currently stands and how this state can be improved. If depression or anxiety is higher at the end of the session (SUDS) or the client is distressed there may not be clarity about how the client might be able to change.

Intensity of the alliance maybe determined by whether the therapist and client are willingly and enthusiastically engaged in task of discussing how the client may change as a result of these sessions/therapy. The client is responsive to the task and the therapist is responsive to the clients needs for suggesting the task. For example, there may be frequent instances of where there is clarity about how the client might be able to change, but if there appears to be little willingness or enthusiasm to take part in this task and the reason for this lack of engagement is not discussed, a lower score may be considered.

5.   There are frequent instances where there is clarity about how the client might be able to change as a result of these sessions/therapy. There is a **VERY HIGH** level of engagement between the therapist and client when they discuss how the client may change.

4.   There are very often instances of clarity about how the client might be able to change as a result of these sessions/therapy. However, there is **ONE** instance where there is concern and this concern is not discussed. There is a **HIGH** level of engagement between the therapist and client when they discuss how the client may change.

3.   There are fairly often instances of clarity about how the client might be able to change as a result of these sessions/therapy. However, there are **TWO** instances where there is concern and this concern is not discussed. There is a **MODERATE** level of engagement between the therapist and client when they discuss how the client may change.

2.   There are sometimes instances of clarity about how the client might be able to change as a result of these sessions/therapy. However, there are **THREE** instances where there is concern and this concern is not discussed. There is a **LOW** level of engagement between the therapist and client when they discuss how the client may change.

1.   There seldom instances of clarity about how the client might be able to change as a result of these sessions/therapy. There are **FOUR OR MORE** instances where there is concern and this concern is not discussed. There is a **VERY LOW, LITTLE OR NO** engagement between the therapist and client when they discuss how the client may change.

8. Task Item

①	②	③	④	⑤
Seldom	**Sometimes**	**Fairly Often**	**Very Often**	**Always**

The client believes that the way they are working with his/her problem is correct.

This item is more reflective of agreement than understanding. Client's feedback is positive in regards to therapy, current intervention or homework assignment, monitoring of mood, and conceptualization. Client perceives the benefits of therapy and has a sense of mastery with acquired skills.

Intensity of the alliance maybe determined by the client's willingness and degree of comfort to reflect on his /her belief that the way they are working with his/her problem is correct and the therapists willingness to discuss these reflections. The client is responsive to the task and the therapist is responsive to the clients needs for suggesting the task. Ruptures may occur in the relationship if there is disagreement about the task. For example, there may be frequent instances of the client reflecting that the way they are working with his/her problemis correct but if there appears to be little willingness or enthusiasm to take part in this task from both therapist and client and the reason for this lack of engagement is not discussed, a lower score may be considered.

5. There are frequent instances of the client believing that the way they are working with his/her problem is correct. There is a **VERY HIGH** level of engagement between the therapist and client when the client reflects his/her beliefs.

4. There are very often instances of the client believing that the way they are working with his/her problem is correct. However, there is **ONE** instance where the client expresses his/her doubts and these doubts are not discussed. There is a **HIGH** level of engagement between the therapist and client when the client reflects his/her beliefs.

3. There are fairly often instances of the client believing that the way they are working with his/her problem is correct. However, there are **TWO** instances where the client expresses his/her doubts and these doubts are not discussed. There is a **MODERATE** level of engagement between the therapist and client when the client reflects his/her beliefs.

2. There are sometimes instances of the client believing that the way they are working with his/her problem is correct. However, there are **THREE** instances where the client expresses his/her doubts and these doubts are not discussed. There is a **LOW** level of engagement between the therapist and client when the client reflects his/her beliefs.

1. There are seldom instances of the client believing that the way they are working with his/her problem is correct. There are **FOUR OR MORE** instances where the client expresses his/her doubts and these doubts are not discussed. There is a **VERY LOW, LITTLE, OR NO** engagement between the therapist and client when the client reflects his/her beliefs.

9. Bond Item

①	②	③	④	⑤
Seldom	**Sometimes**	**Fairly Often**	**Very Often**	**Always**

There is a <u>MUTUAL LIKING</u> between the client and therapist.

Mutual liking involves having pleasant and warm interactions, being engaged, being sensitive, and showing personal interest in each other (e.g., the therapist should enhance or reinforce client expressions of personal values or outside interests, and the therapist should not forget important details about the client's life). There are open expressions of rapport, empathy, warmth, concern, genuineness and humour. Body language is open and relaxed and there are warm tones in the client's and therapist's voice. **In this item the emphasis is on mutual and refers to feelings and experiences between the client and therapist.**

Rating the intensity of mutual liking between the therapist and client can be determined by positive, warm versus negative, cold, hostile interactions, sensitivity, empathy, rapport, open and relaxed versus closed off body language.

5. There are frequent instances of mutual liking between the therapist and client or no explicit need for mutual liking to be present (e.g., if there is no pain expressed by the client). There is an **OPTIMAL** level of warmth, sensitivity, empathy, rapport and open and relaxed body language. There are warm tones in the clients and therapists voices. Any ruptures/strains are immediately investigated and repaired.

4. There are very often instances of mutual liking between the therapist and client. However, there was **ONE** instance when mutual liking may have been appropriate, but was not expressed. There is a **SATISFACTORY** level of warmth, sensitivity, empathy, rapport and open and relaxed body language. Any ruptures/strains are investigated and every effort is made to repair them.

3. There are fairly often instances of mutual liking between the therapist and client. However, there is **TWO** instances when mutual liking was clearly needed, but unexpressed. (e.g., the therapist focuses on the task at hand without being attuned to the clients need for validation). There is a **MODERATE** level of warmth, sensitivity, empathy, rapport and open and relaxed body language. The relationship may seem strained, at times but generally every effort is made to investigate the problem and repair it.

2. There are sometimes instances of mutual liking between the therapist and client. However, there are **THREE** instances where mutual liking was clearly absent (e.g., therapist's voice is mechanical and abrupt). Interactions between therapist and client are **POOR,** but do not seem destructive. Some effort is made to repair ruptures/strain.

1. There are seldom instances of mutual liking between the therapist and client. There are **FOUR OR MORE** instances when the unaddressed need was more blatant and mutual liking didn't happen as often as it should. Interactions are **EXTREMELY POOR** and destructive. Interactions seem negative, cold, and hostile. No effort is made to repair ruptures/strains.

10. Bond Item

①	②	③	④	⑤
Seldom	**Sometimes**	**Fairly Often**	**Very Often**	**Always**

The client and the therapist <u>RESPECT</u> each other.

Respect is indicated by both therapist and client placing a high value on the other's contribution. It is also indicated by therapist sensitivity to client problems and appropriate boundaries being set and kept. Respect depreciates when the client questions the therapist's competence/ability to help, when either the therapist or client interrupts the other with irritation or impatience, or when the therapist blames or sees the client as incompetent in dealing with his or her problems. No response can lead to the client feeling unimportant, discredited or doubtful. On the other hand pauses are important and silences may indicate that the therapist or client need to reflect on what the other has said before they speak.

Rating the intensity of respect between the therapist and client can be determined by sensitivity and appropriate boundaries being set and kept.

5. There are frequent instances of respect between the therapist and client or no explicit need for respect between the therapist and client to be present. There is an **OPTIMAL** level of respect. Appropriate boundaries are set and kept and there is no difficulty for the therapist to convey confidence and competence. Any ruptures/strains are immediately investigated and repaired.

4. There are very often instances of respect between the therapist and client. However, there was **ONE** instance when respect may have been appropriate, but was not expressed. There is a **SATISFACTORY** level of respect. There are no significant interpersonal problems regarding respect between the therapist and client. Any ruptures/strains are investigated and every effort is made to repair them.

3. There are fairly often instances of respect between the therapist and client. However, there is **TWO** instances when respect was clearly needed, but unexpressed. There is a **MODERATE** level of respect. The level of respect between the therapist and client may seem strained, but generally every effort is made to investigate the problem and repair it.

2. There are sometimes instances of respect between the therapist and client. However, there are **THREE** instances where respect was clearly absent. Interactions between therapist and client are **POOR**, but do not seem destructive. The therapist or client tend to interrupt each other with irritation or impatience, but some effort is made to repair ruptures/strains.

1. There are seldom instances of respect between the therapist and client. There are Four OR MORE instances when the unaddressed need was more blatant and respect didn't happen as often as it should. Interactions between therapist and client are **EXTREMELY POOR,** and destructive. The therapist or client interrupt each other with irritation or impatience, or the therapist blames or sees the client as incompetent in dealing with his or her problems. The relationship is negative and may seem hostile. No effort is made to repair ruptures/strains.

11. Bond Item

①	②	③	④	⑤
Seldom	**Sometimes**	**Fairly Often**	**Very Often**	**Always**

The <u>CLIENT FEELS THE THERAPIST RESPECTS AND CARES ABOUT THE CLIENT, EVEN WHEN THE CLIENT DOES THINGS THAT THE THERAPIST DOES NOT APPROVE OF</u>.

The client is not concerned about being judged negatively by the therapist for something he or she says or does. There are open expressions of "unconditional positive regard" (e.g., client feels free to express opinions or behaviors that may unethical or concerning, therapist does not use humor or jokes about the client, client does not appear to be fearful of rejection or being judged).

Rating the intensity of unconditional positive regard between the therapist and client can be determined by the client perception of not being judged negatively by the therapist for something he/she says or does.

5. There are frequent instances of the client feeling that the therapist respects and cares about the client, even when the client does things that the therapist does not approve of or no explicit need for unconditional positive regard to be present. There is an **OPTIMAL** unconditional positive regard between the therapist and client. Any ruptures/strains are immediately investigated and repaired.

4. There are very often instances of the client feeling that the therapist respects and cares about the client, even when the client does things that the therapist does not approve of. However, there was **ONE** instance when unconditional positive regard may have been appropriate, but was not expressed. There is a **SATISFACTORY** level of unconditional regard. There are no significant interpersonal problems. Any ruptures/strains are investigated and every effort is made to repair them.

3. There are fairly often instances of the client feeling that the therapist respects and cares about the client, even when the client does things that the therapist does not approve of. However, there is **TWO** instances when unconditional positive regard was clearly needed, but unexpressed. There is a **MODERATE** level of unconditional positive regard. The relationship between the therapist and client seems strained and but every effort is made to investigate the rupture/strain and repair it.

2. There are sometimes instances of the client feeling that the therapist respects and cares about the client, even when the client does things that the therapist does not approve of. However, there are **THREE** instances where unconditional positive regard was Interactions between therapist and client are **POOR**. The client does not openly express himself/herself and appears guarded. Some effort is made to repair ruptures/strains.

1. There are seldom instances of the client feeling that the therapist respects and cares about the client, even when the client does things that the therapist does not approve of. There are **FOUR OR MORE** instances when the unaddressed need was more blatant Interactions between therapist and client are **EXREMELY POOR** and are destructive. The relationship seems negative and hostile. No effort is made to repair ruptures/strains.

12. Bond Item

The client feels that the therapist <u>APPRECIATES HIM/HER AS A PERSON.</u>

Appreciation involves expressions by the therapist of non-judgmental acceptance, warmth, empathy, personal interest, and **sensitivity to the uniqueness of the client's plight**. **However, this item refers to the clients perceptions.** The client is willing to talk about what is helpful in therapy. For example, client offers agenda items, talks about their week's experiences and learning during review of homework and update at the onset of the session. Client offers feedback on how therapy can be improved during session feedback, or client is willing to talk about coping strategies and beliefs. Does the therapist provide positive reinforcement, praise, and encouragement. Are there missed opportunities to appreciate the clients progress.

Rating the intensity of the client feeling that the therapist appreciates him/her as a person can be determined by therapist of non-judgmental acceptance, warmth, empathy, personal interest, and sensitivity to the uniqueness of the client's plight.

5. There are frequent instances of the client feeling that the therapist appreciates him/her as a person or no explicit need for is the client to feel that the therapist appreciates him/her as a person to be present. There is an **OPTIMAL** level of therapist non-judgmental acceptance, warmth, empathy, personal interest, and sensitivity to the uniqueness of the client's plight. Any ruptures/strains are immediately investigated and repaired.

4. There are very often instances of the client feeling that the therapist appreciates him/her as a person. However, there was **ONE** instance when appreciation may have been appropriate, but was not expressed. There is a **SATISFACTORY** level of therapist of non-judgmental acceptance, warmth, empathy, personal interest, and sensitivity to the uniqueness of the client's plight. There are no significant interpersonal problems regarding the client feeling that the therapist appreciates him/her as a person. Any ruptures/strains are investigated and every effort is made to repair them.

3. There are fairly often instances of the client feeling that the therapist appreciates him/her as a person. However, there is **TWO** instances when appreciation was clearly needed, but unexpressed. There is a **MODERATE** level of therapist appreciation for the client as a person. The relationship seems strained, but generally every effort is made to investigate the problem and repair it.

2. There are sometimes instances of the client feeling that the therapist appreciates him/her as a person. However, there are **THREE** instances where appreciation was clearly absent. Interactions between therapist and client are **POOR,** but do not seem destructive. Some effort is made to repair ruptures/strain.

1. There are seldom instances of the client feeling that the therapist appreciates him/her as a person. There are **FOUR OR MORE** instances when the unaddressed need was more blatant and appreciation didn't happen as often as it should. Interactions are **EXTREMELY POOR** and destructive. There is sensitivity to the uniqueness of the client's plight. The relationship seems negative and hostile. No effort is made to repair ruptures/strains.

Appendix D

Confidentiality Form for Raters

Center for Psychology, School of Psychology, Massey University

CONFIDENTIALITY FORM

We maintain a strict and firm policy of confidentiality on all Centre matters and absolutely no information about clients will be passed on to another person or agency without the client's written consent. The two main exceptions to this policy are a life threatening emergency or court order of records (a request from a lawyer is not an exception). Case materials may be used for teaching purposes or research, but only under strict assurance that identifying information will not be included in any such presentation unless the client gives written permission.

Maintaining confidentiality for client attending the Centre is a high priority. Discussing confidential case material outside of the context of the Centre setting is a breach of confidentiality. Any discussion of case material should be limited to the duties which bring you into contact with this material. If you are a student in training or research assistant it is expected that you will discuss case or research material with the responsible clinician who is supervising your involvement in a particular case. You may discuss this material freely with your supervisor or the Director of the Centre of Psychology. Case discussion in the context of clinical training within the confines of the Centre building is permitted, but the identity of the client should remain confidential. If you have any further questions about the issue of confidentiality you consult with your supervisor.Take care not to discuss cases or use client names where you may be overheard by others. This is particularly important in the reception area when you may be using a telephone or speaking with the secretary. Often clients are waiting to be seen and it is disconcerting to hear others names who may or may not be clients.

You are not permitted to acknowledge that a client has been seen at the Centre unless you have written permission from the client. This is particularly important when responding to enquiries over the telephone. Even if you have permission you should be absolutely certain of the identity of the caller before providing information. Legitimate callers fully understand the issue of confidentiality and accept that you can not discuss a case under such circumstances. You may wish to have the caller hang-up and contact them at a verifiable number (e.g., ACC office) before discussing the case.

If you are writing reports using computer word-processing, you should be aware that most word-processing programmes automatically make back-ups of the file. You should use a secure computer wherever possible. If you must leave a report on a computer disk or hard drive, the computer or disk should be locked in a room or file drawer, access to the computer hard drive or disk should be restricted to only those who are involved in a case, and the keyboard should be locked when the computer is not in use. You should never leave client information unattended on screen. You are encouraged to leave client identifying information out of a partially completed report and enter this last. You might use the word "client" instead of the actual name and at the completion of the report use the Search and Replace functions to enter the client name. After the report is completed and you have entered any relevant information into a Database you should delete all files (including backup files from your disk or hard drive).

If you print a report using the Center printer, be aware that this is a shared printer and you should collect the report *immediately* so as to avoid leaving it unattended or forgetting it.Finally, all records should be filed in a locked file cabinet at the end of the day. Do not leave records or other confidential information unattended on your desk.

I have read the Massey University Centre for Psychology Confidentiality Form and have had an opportunity to have my questions regarding this policy answered by the Centre Director. I agree to maintain client confidentiality as outlined in the Centre for Psychology Confidentiality Form.

Name: Date:
Signed: Date:
Witness: Date:

Appendix E

Training Package for Rating the WAI-SR-O

Alliance and Syptomatic Change within Cognitive-Behaviour Therapy for Depression

Training Package for Rating the WAI-SR-O

Carol Osborne
Nikolaos Kazantzis, Ph.D

February 2008

First day of Training

- Brief overview of my research

- Practice Exercise: What we already know about the alliance.

- Alliance within CBT

- The Working Alliance Inventory (WAI)

- Overview of Bordons theory (Goal/Task/Bond)

- Rating the WAI-SR-O

- Practical Exercise: Rating Session One

Overview of Research

- Important to define the term alliance.

- Important to define alliance within CBT.

- Find the most appropriate instrument to operationalize the alliance concept within CBT.

Practical Exercise: What we already know about the Alliance

- An exercise to activate your existing knowledge!

- Collaboratively draw a mind map

- Everything you know about the therapist-client relationship or alliance

- Use others ideas to stimulate new ideas

- Use colour

- Use pictures

- Be creative

- Have fun

Alliance

- Broadly defined as the collaborative bond between the therapist and client (Krupnick et al., 1996).

- Essential ingredient of any psychotherapy (Gaston, 1990).

- Most theoretical definitions of the alliance have three themes in common:
  (a) the collaborative nature of the relationship,
  (b) the affective bond between the client and the therapist, and
  (c) the client's and therapist's ability to agree and work on treatment goals and tasks (Bordin, 1979;Gaston, 1990; Horvath & Luborsky, 1993;

Alliance within CBT

- Alliance serves as a positive background for the actual work of therapy to be carried out in.

- Aaron Beck emphasized that therapist characteristics, such as warmth, empathy, respect, genuineness and confidence were necessary in therapy to engage their clients in a process of 'collaborative empiricism' (Beck et al., 1979). However, he also believed that these non-specific factors were not sufficient to produce optimum therapeutic effect.

- Collaborative empiricism encapsulates the idea that the client and therapist work as a team.

Collaborative Empiricism

- Collaboration is characterised by mutuality and active negotiation between therapist and client.

- It ensures that the client and the therapist have compatible goals at each point in the course of therapy.

- Prevents misunderstandings

- Minimizes ruptures within the relationship.

Collaborative Empiricism

- Establishing an agenda in a collaborative manner to structure the therapy session.

- Bridging checks on client's perception and understanding of the previous session.

- Providing clear rationale for therapy process/ interventions/homework assignments.

- Socratic questioning and guided discovery encourages the client to openly explore and discover their thoughts, feelings and behaviour for themselves.

Collaborative Empiricism

- Listening, reflecting, paraphrasing and summarizing regularly.

- A bond of trust and positive regard is established and a secure background created in which it was safe to discuss difficult situations, including ruptures or strains in the alliance.

- Eliciting feedback and providing feedback at the end of each session.

Alliance Measurement

- It is important to select an appropriate alliance measure to fit the principles of CBT.

- Six of the main alliance instruments (Penn, VTAS, CALPAS & 3 versions of the original WAI) were compared and found to all have acceptable levels of internal consistency (i.e., alphas were above .90), reliability, and inter-rater reliability (i.e., intraclass correlation coefficients were all .7 or above) (see Cecero, Fenton, Nich, Frankforter, & Carroll 2001; Tichenor & Hill,1989)

The Working Alliance Inventory

- Widely used.

- Most transtheoretical and fits CBT principles.

- Based on Bordin's theory of alliance (Bordin, 1979, 1994).

- The original WAI of 36 items (Horvath & Greenberg, 1989) consists of three subscales (i.e., Bond, Tasks and Goals).

Summary of WAI Versions

- <u>Original WAI (36 items)</u>
- Client Version (Horvath & Greenberg, 1989)
- Therapist Version (Horvath & Greenberg, 1989)
- Observer Version (Tichenor & Hill, 1989).

- <u>Shortened WAI (12 items)</u>
- Client Version (Tracey & Kokotovic, 1989).
- Therapist Version (Tracey & Kokotovic, 1989).
- Observer Version (Andrusyna et al., 2001).

- <u>Shortened Revised WAI (12 items)</u>
- Client Version   (Hatcher & Gillaspy, 2006)
- Observer Version (Client version modified for the current research)

Working Alliance Inventory-Short Form–Revised–Observer Version (WAI-SR-O)

- The WAI-SR-O (handout of scale) consists of three subscales (Goals, Tasks and Bond) (Hatcher & Gillaspy, 2006).

Working Alliance Inventory–Short Form-Revised- Observer Version (WAI-SR-O)

Goal Items

1. The client and therapist collaborated on setting the goals for the session.

2. There is agreement on what is important for the client to work on.

3. The client and therapist are working on mutually agreed upon goals.

4. The client and therapist have established a good understanding of the changes that would be good for the client.

Working Alliance Inventory-Short Form–Revised–Observer Version (WAI-SR-O)

- **Goals** are defined as the objectives of the therapist and client that specifically target change. Goals can include:
- Reduction in symptoms,
- Improvement in interpersonal skills or relationships,
- Increase in self-esteem,
- Development of new ways of thinking or behaving.

Working Alliance Inventory-Short Form–Revised–Observer Version (WAI-SR-O)

- Agreement is assessed according to the extent to which both therapist and client see the goals as important, clear and capable of being accomplished.

- Understanding of the benefits resulting from the change.

Working Alliance Inventory-Short Form–Revised–Observer Version (WAI-SR-O)

Task Items

5. There is agreement about the usefulness of the current activity in therapy (i.e., the client is seeing new ways to look at his/her problem)

6. There is agreement that what the client and therapist are doing in therapy will help the client to accomplish the changes he/she wants.

7. As a result of these sessions there is clarity about how the client might be able to change.

8. The client believes that the way they are working with his/her problem is correct.

Working Alliance Inventory-Short Form–Revised–Observer Version (WAI-SR-O)

- **Tasks** are defined as therapeutic techniques or processes which help clients to increase awareness of their own thoughts, feelings, values and needs.

  Examples of therapeutic tasks include:
- support/reassurance
- reflection
- reformulation
- cognitive restructuring

Working Alliance Inventory-Short Form–Revised–Observer Version (WAI-SR-O)

Ratings on tasks are made according to:

- How responsive the client is to the task.

- How responsive the therapist is to client need in suggesting the task.

- The tasks should be seen as important, appropriate, and clear by both the client and the therapist.

Working Alliance Inventory-Short Form–Revised–Observer Version (WAI-SR-O)

Bond Items

9. There is a mutual liking between the client and therapist.

10. The client and the therapist respect each other.

11. The client feels that the therapist appreciates him/her as a person.

12. The client feels the therapist respects and cares about the client, even when the client does things that the therapist does not approve of.

Working Alliance Inventory-Short Form–Revised–Observer Version (WAI-SR-O)

- **Bond** is defined as the mutual liking, respect, and trust between the client and the therapist.
- Characterised by therapist genuineness, warmth, and understanding.
- Client confidence in the therapist, client comfort and reciprocal respect are also included.
- Bond is assessed through tone of voice, amount and quality (i.e., degree of comfort) of client talk concerning intimate issues.
- Therapist degree of comfort, therapist non-defensiveness, therapist accurate empathy (especially validated by the client) and the mutual value placed on each others contributions.

Rating Guidelines

- Ratings are based on frequency of occurrence within the session. However, intensity should also be taken into account.

- Most important criteria is the clients agreement with the therapist, belief in the therapists care and competence, and acceptance and participation in strategies used (e.g., willingly discloses, evaluates cognitions or tries new behaviors).

Rating Guidelines

- Ratings should be based on subjective reaction and observable behavior, that is, clinical judgment.

- Ratings also need to be based on what is going on between the therapist and the client (do they seem to be working well together, have a good rapport, reciprocal relationship, the client is working and getting value out of the sessions). Care needs to taken that ratings are not based on the therapists style of relating.

Rating Guidelines

- Consider all relevant information from a session. Try not to be distracted by elements that are particularly salient. For example, the client may disagree on some task or goal, yet come to an agreement later in the course of the session. Thus, take into account the information from the entire session.

- Do not consider the pathology or other characteristics of the client when making ratings. Raters should refer to realistic prototypes of good and poor alliances and not be idiosyncratic.

Rating Guidelines

- Do not rate the technical skill or competence of the therapist exhibited in the session. Although they may be related to competence, these ratings concern the therapeutic alliance. This rating scale places greater emphasis on timing, warmth, style, respect of client needs, than application of techniques or interventions.

Rating Guidelines

- Ratings on this form assume the relationship quality is present. Subtract ratings when a therapist alliance behavior would have been helpful, when a therapist alliance behavior was clearly unhelpful, or when a therapist alliance behavior was absent. Simultaneously, the intensity or strength of the affective bond and the ability of the therapist and the client to work together in therapy needs to be taken into consideration. Thus, start with point '5' on each rating scale, and move down the scale as more therapeutic alliance elements are missing.

Guidelines for Scoring Specific Items

1. Goal Item
The client and therapist COLLABORATED ON SETTING THE GOALS for the session.

- Both therapist and client work as a team to set the content of the in-session discussion or agenda.
- Collaboration is assessed by the responsiveness of the therapist and client to each other's initiations.
- Points should be deducted if one
- participant changes the topic and the other continues to bring it.
- Consideration needs to be given to the therapist (a) politely interrupting and trying to bring client back on track, (b) discussing the client's reason for not wanting to keep to the agenda, or (c) re-negotiating session agenda.

Guidelines for Scoring Specific Items

2. Goal Item
   There is <u>AGREEMENT ON WHAT IS IMPORTANT</u> for the client to work on.

- Goals may refer to topics and concerns (in-session and long term goals) raised by the therapist and the client.
- "Important" means therapeutically significant for the client.
- Agreement could include determining which problems are the most important to work on during the session/therapy, and the prioritizing of long term therapy goals and in-session agenda
- Therapist monitors flow of discussion.

Guidelines for Scoring Specific Items

3. Goal Item
The client and therapist are <u>WORKING ON MUTUALLY AGREED UPON GOALS</u>

- This item refers more to in-session goals or agenda items.

- For example, the client and therapist work on agreed prioritized agenda items, therapist and client participate in 'in-session' activity which is linked to agenda/therapy goals, summaries are provided, and/or feedback and bridge provides transition between topics.

Guidelines for Scoring Specific Items

4. Goal Item
The client and therapist have established a <u>GOOD UNDERSTANDING OF THE CHANGES THAT WOULD BE GOOD FOR THE CLIENT</u>

- Change implies that there is some degree of a focus on the future. The focus here should be on the client's goals for therapy and what they might actually do (the specific content) to reach these goals. For example, the client and therapist collaboratively agree on therapy goals, therapy goals are prioritized, a rationale is provided for interventions/homework assignments and this rationale is linked back to the client's therapy goals).

Guidelines for Scoring Specific Items

5. Task Item
  There is agreement about the <u>USEFULNESS OF THE CURRENT ACTIVITY</u> in therapy (i.e., the client is seeing new ways to look at his/her problem).

- Activity is restricted to that within the session.

- Agreement about the usefulness of the current activity in therapy could include: agreement between client and therapist on about the rationale, importance, or usefulness of an intervention or homework assignment, and/or generalization of learning.

Guidelines for Scoring Specific Items

6. Task Item
There is <u>AGREEMENT</u> that what the client and therapist are doing in therapy will help the client to accomplish the changes he/she wants.

- Agreement about chosen interventions and homework assignments.
- Interventions and homework assignments are related to client goals.
- The homework assignment is important to the client and they are ready and confident to complete it.
- Barriers to complete an intervention or homework assignment are discussed.

Guidelines for Scoring Specific Items

7. Task Item
As a result of these sessions there is clarity about how the client might be able to <u>CHANGE.</u>

- Change implies that there is some degree of a focus on the future. The focus here should be on how the client might change (the process of change/ways of dealing with problems) not on what the client might actually do (the specific content).

Guidelines for Scoring Specific Items

8. Task Item
The client believes that the way they are working with his/her problem is correct.

- This item is more reflective of agreement than understanding. Client's feedback is positive in regards to therapy, current intervention or homework assignment, monitoring of mood, and conceptualization. Client perceives the benefits of therapy and has a sense of mastery with acquired skills.

Guidelines for Scoring Specific Items

9. Bond Item
There is a <u>MUTUAL LIKING</u> between the client and therapist.

- Pleasant and warm interactions,
- Sensitivity,
- Showing personal interest in each other
- Open expressions of rapport, empathy, warmth, concern, genuineness and humour.
- Body language is open and relaxed
- Warm tones in the client's and therapist's voice.

Guidelines for Scoring Specific Items

10. Bond Item
The client and the therapist <u>RESPECT</u> each other.

- Both therapist and client place high value on the other's contribution.
- Therapist sensitivity to client problems
- Appropriate boundaries being set and kept.
- Respect depreciates when the client questions the therapist's competence/ability to help, when either the therapist or client interrupts the other with irritation or impatience, or when the therapist blames or sees the client as incompetent in dealing with his or her problems.

Guidelines for Scoring Specific Items

11. Bond Item
The <u>CLIENT FEELS THE THERAPIST RESPECTS AND CARES ABOUT THE CLIENT, EVEN WHEN THE CLIENT DOES THINGS THAT THE THERAPIST DOES NOT APPROVE OF</u>.

- The client is not concerned about being judged negatively by the therapist for something he or she says or does.

- There are open expressions of "unconditional positive regard" (e.g., client feels free to express opinions or behaviors that may be unethical or concerning, therapist does not use humor or jokes about the client, client does not appear to be fearful of rejection or being judged).

Guidelines for Scoring Specific Items

12. Bond Item
The client feels that the therapist APPRECIATES HIM/HER AS A PERSON

- Therapist expressions of non-judgmental acceptance, warmth, empathy, personal interest, and sensitivity to the uniqueness of the client's plight.

- The client is willing to talk about what is helpful in therapy.

Appendix F

Case Results

Case 1 ("Jayde")

Background Information

Jayde was a 25 year old female of Pacific Islander and European descent. She had a partner, but no children. She was employed full time as a personal assistant and occasionally she worked weekends as a retail assistant. Jayde presented as slim and well-groomed and was easily engaged in therapy. She attended therapy sessions in her lunch hour. Jayde tended to be very conscious of her appearance and her weight.

She presented with symptoms of sadness, anger, irrational behaviour ("something takes me over"), "edgy", lack of motivation, lost of interest in work, lack of concentration, and easily distracted, late insomnia, increased appetite for sweet foods and difficulty socializing with friends. Suicidal thoughts "come and go", but she had not experienced these thoughts recently. She would "not do anything. It is not an option". At 23 years she was diagnosed with anxiety by her General Practitioner who prescribed Aropax for six months. Symptoms at this time included dizziness, lack of concentration, tunnel vision, tingling skin, and loss of confidence.

Jayde enjoyed playing netball and had surgery on her knee due to a netball injury. Two weeks before the initial assessment her surgeon advised another operation. Jayde reported that this advice had "knocked her" and she perceived that her indecision about the surgery has precipitated her depressive symptoms. She sought therapy because she did not want to go onto medication again. However, she was also skeptical that therapy would help her.

At the initial assessment Jayde was diagnosed as having a major depressive episode and she scored 24 on the BDI-II which placed her depression severity at a moderate level. The CIDI also assessed Jayde as also having panic disorder without agoraphobia.

Results

Depression Severity

At the pretreatment assessment Jayde's depression severity was 24 points on the BDI-II. Her depression severity increased to 28 points in S 1. During the first week of therapy the severity of Jayde's depressive symptoms decreased. Her BDI-II score dropped 7 points in S 2. However, the severity of her depressive symptoms increased the second week of therapy and continued to gradually increase over the third week. There was a slight drop in S 5, but in S 6 Jayde rated herself 27 on the BDI-II. Her depression severity decreased 2 points in S 7 and then a further 8 points from in S 8. This score was maintained until S10, where it rose 1 point to 18. Jayde continued therapy until session 20 and her final BDI-II score was 16 (ie., mild depression). At the 2 monthly follow up Jayde's BDI-II score was 13. Jayde had just broken up with her partner at the 6 monthly follow-up and her score on the BDI –II had increased to 22. At the end of her therapy Jayde was offered other treatment options.

Graph displaying Jayde's Depression Severity scores

Depression Severity
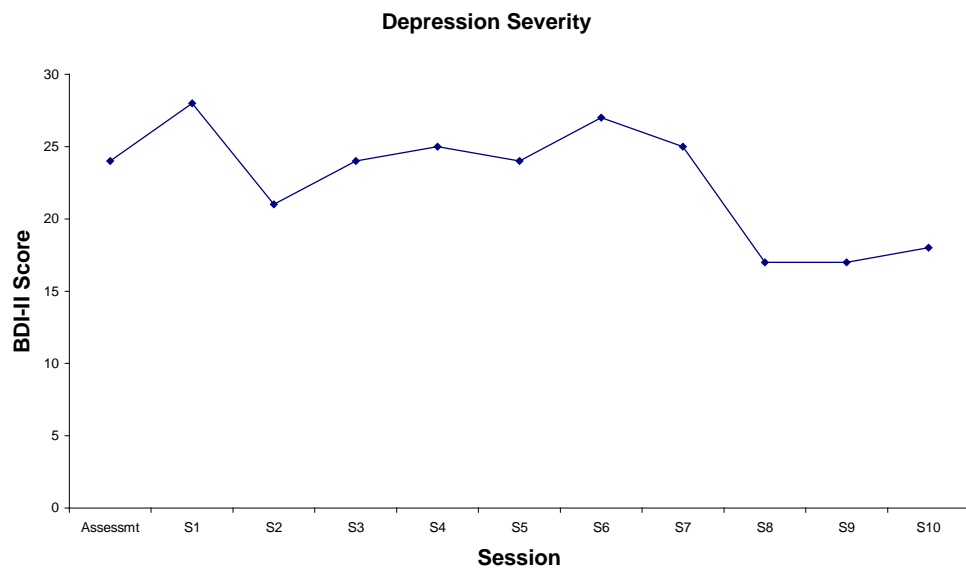
Alliance

Total Alliance

The Total Alliance score fluctuated each session over the first four sessions between the scores of 41 and 53 on the WAI-SR-O. The Total Alliance score increased in the first week of therapy, decreased in S 3 and then increased again in S 4. It then remained fairly stable with slight 1-2 point fluctuations over the next five sessions before decreasing in S 10. Throughout the ten sessions the Total Alliance score fell into the high range of the WAI-SR-O with scores ranging between 41 and 53.

Graph showing Total Alliance Scores for Jayde.

Total Alliance
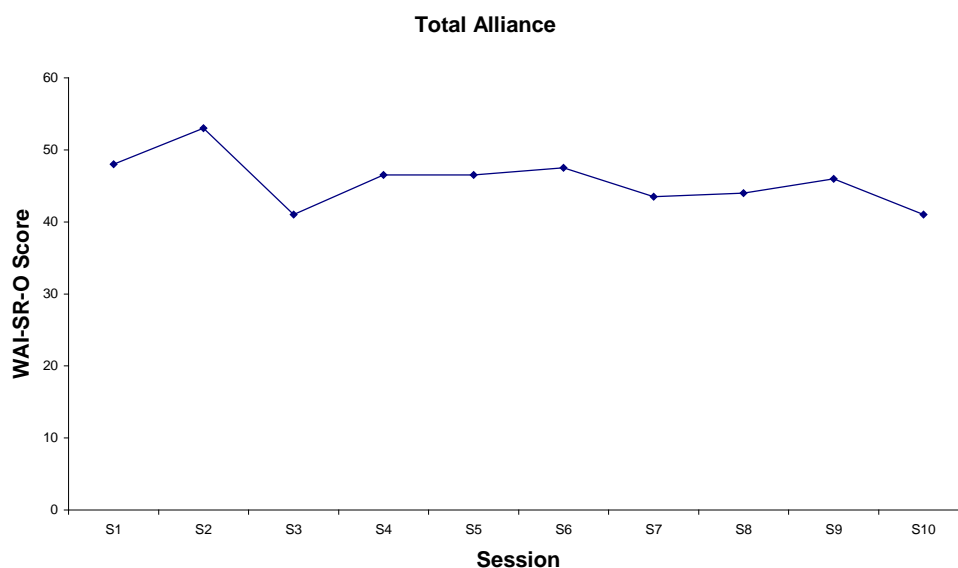
Goal Subscale

The Goal subscale score fluctuated over the ten sessions. It increased over the first week of therapy to a score of 18, decreased in S 3 and then gradually increased again until S 6, before dropping sharply in S 7. The Goal subscale score then gradually increased again to S 9, but decreased in S 10. The Goal dimension of the alliance was the most developed in S 5 only.
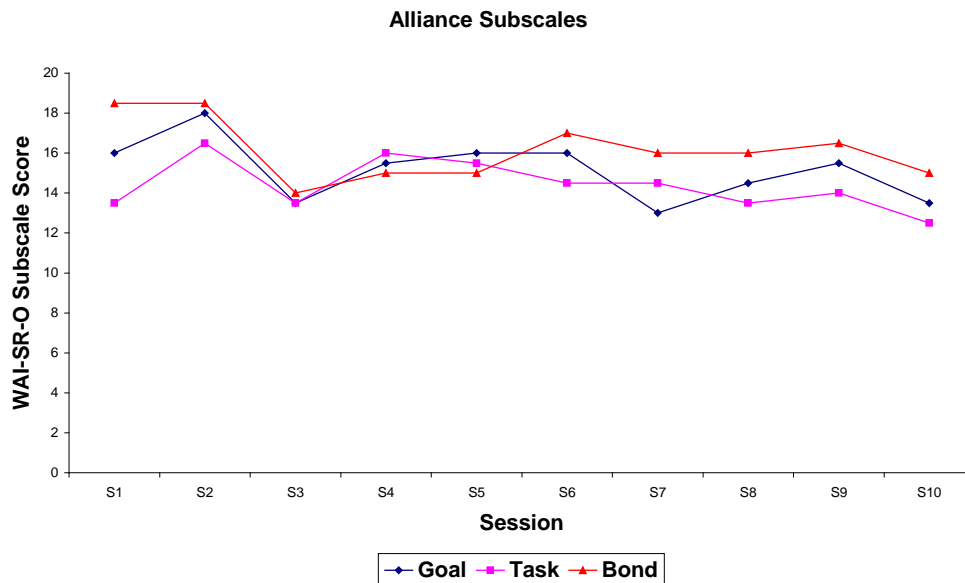
Task Subscale

The Task subscale score fluctuated weekly over the ten sessions. It increased in S 2, decreased in S 3, increased again in S 4 and then gradually decreased for the remainder of the sessions. The Task dimension of the alliance was the most developed in S 4 only.

The score of the Bond subscale maintained over the first two sessions at a high score of 18.5, but it decreased in S 3 before gradually increasing to S 6. Bond subscale scores were fairly stable over the latter third of the ten sessions, but it decreased at S 10. The Bond dimension of the alliance was the most developed throughout most of the sessions.

Graph displaying Goal, Task and Bond Subscale Scores for Jayde

Alliance and Symptomatic Change

In the first week of therapy there was an increase in the Total Alliance and Goal and Task subscale scores. The Bond subscale score was maintained on a high score of 18.5. During that time period the severity of Jayde's depressive symptoms decreased. There was a decrease in the Total Alliance score and all the subscale scores in the second week and an increase in Jayde's depressive symptoms over that week. When the severity of Jayde's depressive symptoms dropped 2 points on the BDI-II in S 7 and then 8 points in S 8 the Bond subscale score had gradually been increasing from S 4. The Total Alliance and Goal subscale score increased from S 7 to S 9.

Case 2 ("Malcolm")

Background Information

Malcolm was a 62 year old English male. He has belonged to an Eastern religion since he was 21years old. He was married with two adult daughters. He was an experienced retail manager and was currently looking for work. Malcolm was tall, slim and dressed tidy casual. He was easy to engage in therapy, but would often go off topic.

His presenting symptoms included: sadness, anhedonia (life was not balanced and he had no freedom or pleasure in his life), irritability/grumpiness (he has started swearing under his breath), feelings of worthlessness and hopelessness, decision making was difficult and he didn't enjoy socializing. While Malcolm's weight was consistent, he craved sweet foods and then he binged on them. Lately he has been more tired and needed more sleep. He has also been questioning his purpose in life. He has had suicidal thoughts (no plan), but this has been difficult for him because he shouldn't have these thoughts at all. The religious faith he belonged to did not believe in suicide. Participating in therapy was also difficult because "if you have a religion you shouldn't need this help".

At the initial assessment Malcolm was diagnosed as having a major depressive episode. He scored 15 on the BDI-II which placed his depression severity at a mild level, but it was noticeable that Malcolm tended to minimize his problems.

Results

Depression Severity

At the pretreatment assessment Malcolm's depression severity was 15 points on the BDI-II. His depression severity decreased to 12 points in S 1. Malcolm's depression severity appeared to fluctuate weekly throughout therapy. In the first week of therapy Malcolm's BDI-II score continued to gradually decrease to 10 in S 3. It increased in S 4 to 11 and this score was maintained throughout the third week of therapy. It decreased to a score of 9 in S 6. From S 7 to S 9 this score fluctuated between 10 and 12, dropping back down to 9 again in S 10. Malcolm finished therapy in Session 13, because he had found work and did not have the time to attend sessions. His final BDI-II score was 1 (i.e., no depression).

Graph displaying Malcolm's Depression Severity Scores
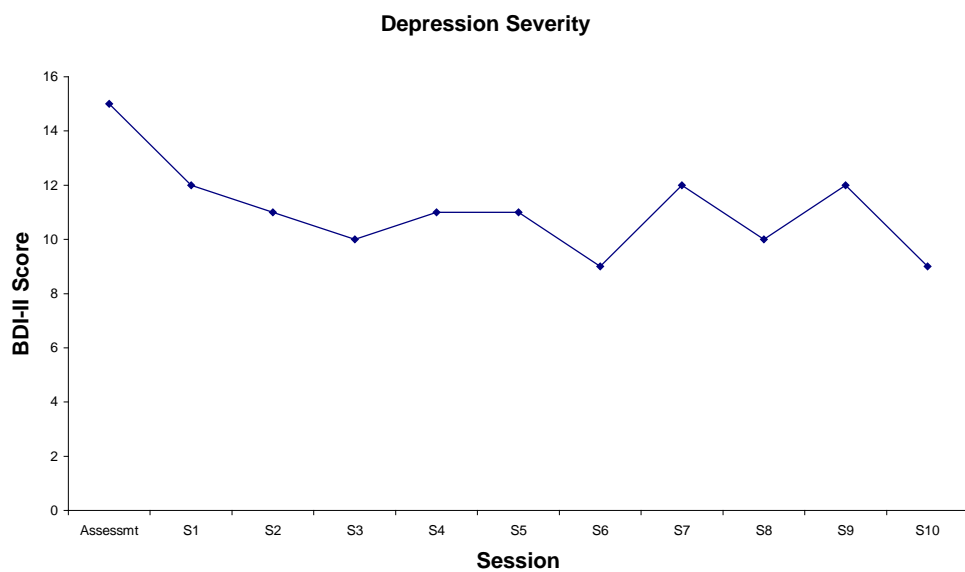
Alliance

<u>Total Alliance</u>

The Total Alliance scores fluctuated throughout the ten sessions. They increased over the first week of therapy, decreased in the second week, and then increased again in the third week. In S 8 the Total Alliance scores decreased before gradually increasing again to S 10. Throughout the ten sessions the Total Alliance score stayed in the high range of the WAI-SR-O with scores ranging between 42 and 50.5

Graph showing Total Alliance Scores for Malcolm.

Total Alliance

WAI-SR-O Score vs *Session* (S1–S10)
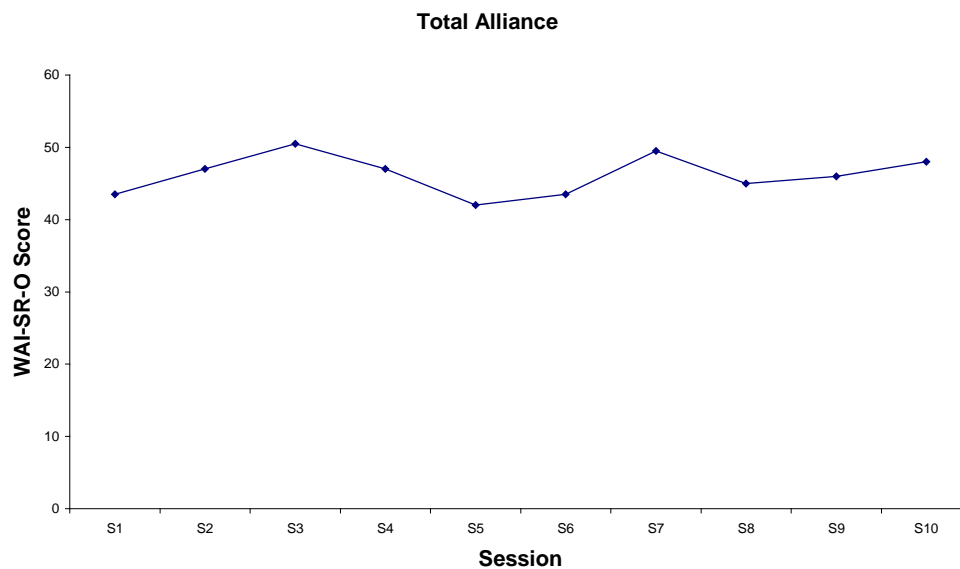
<u>Goal Subscale</u>

The Goal subscale score increased in S 3 and then gradually declined across the remainder of the sessions.

<u>Task Subscale</u>

The Task subscale score increased in the first week, decreased in the second week, increased again in the third and then remained stable for the remainder of the sessions.

252

The Bond dimension of the alliance was the most developed throughout the 10 sessions. At S 2 the Bond subscale score was 19 out of a possible 20. From S 3 to S 5 the Bond subscale score gradually decreased, it then gradually increased to S 7, before fluctuating each session in a fairly stable line until S 10.

Graph displaying Goal, Task and Bond Subscale Scores for Malcolm.

Alliance Subscales
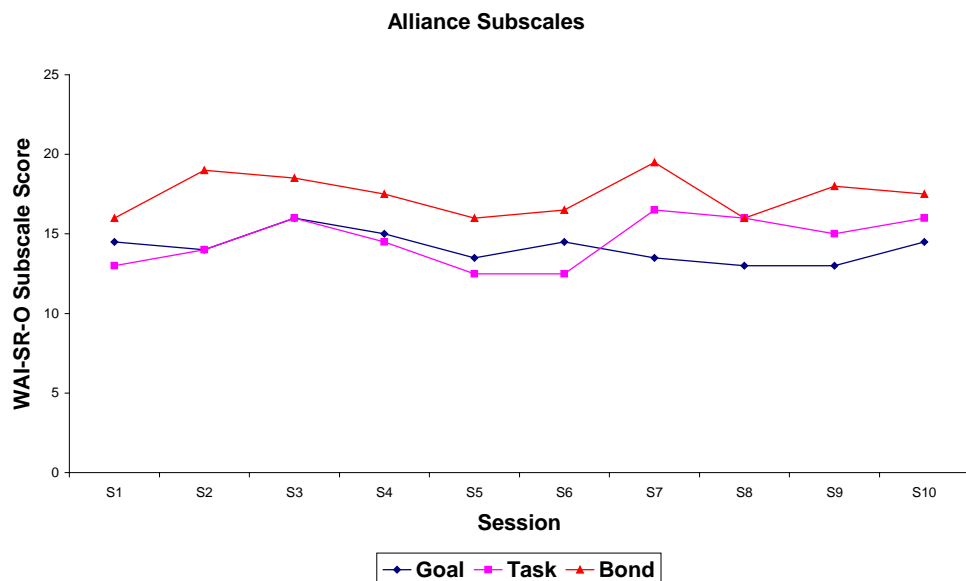
Alliance and Symptomatic Change

In the first three sessions of therapy there was a gradual increase in the Total Alliance score and a gradual decrease in Malcolm's depression severity. From S3 to S 5 there was a gradual decrease in Total Alliance score and an increase in Malcolm's depression severity. Whenever there was a decrease in depression severity in the fourth and fifth week of therapy there was an increase in the Total Alliance and Bond subscale scores in the same session. Furthermore, whenever there was an increase in depression severity there was a decrease in the Total Alliance and Bond subscale scores in the same session.

Case 3 ("Karen")

Background Information

Karen was a 38 year old Caucasian female. She was of average height and build and was dressed tidy casual or in gym gear if she had been to the gym beforehand or was on her way there after therapy. As well as working out in the gym she also enjoyed playing competition softball. She had a sense of humour and was easy to engage in therapy. She had had previous counselling. She was a very expressive person who talked a lot and at times was difficult to interrupt or goal direct.

Karen had been separated one year from her husband after six years of marriage and was a solo mother of a 4 year 6 month old boy and eighteen month old girl. She and her husband tried reconciliation, but Karen ended the relationship 2-3 months before coming to therapy.

Her presenting symptoms included: sadness, tearfulness, feelings of powerlessness and of not being in control of her life, loss of interest in sex, restlessness, agitated, anxiety, loss of appetite with loss of weight and difficulty to concentrate on her studies. She was lonely, discontented and angry ("stuff the world"). She reported that she did not have any suicidal thoughts.

At initial assessment Karen was diagnosed with a major depressive episode and she scored 38 on the BDI-II which placed her depression severity in the severe range. The CIDI assessed Karen as also suffering from bulimia nervosa (from 11 years old), post traumatic stress disorder (from 8 years old), generalized anxiety disorder and hypochondrias.

Results

Depression Severity

At the pretreatment assessment Karen's depression severity was 38 points on the BDI-II. Her depression severity decreased to 32 points in S 1. The severity of Karen's depressive symptoms fluctuated dramatically, but gradually decreased over the ten sessions. Her depressive severity decreased 6 points during the first week of therapy and then increased 5 points in S 4. In the third week of therapy Karen's depression severity decreased 24 points to a score of 11 on the BDI-II in S 5. However, it rose again in S 6 to 22. This score was maintained in S7, but then it decreased 20 points to 2 in S 8. Her depression severity gradually increased from S 9 to S 10. Karen continued therapy until Session 19, only dropping out because she shifted out of the area. Her final BDI-II score was 3 (i.e., no depression).

Graph displaying Karen's Depression Severity Scores

Depression Severity
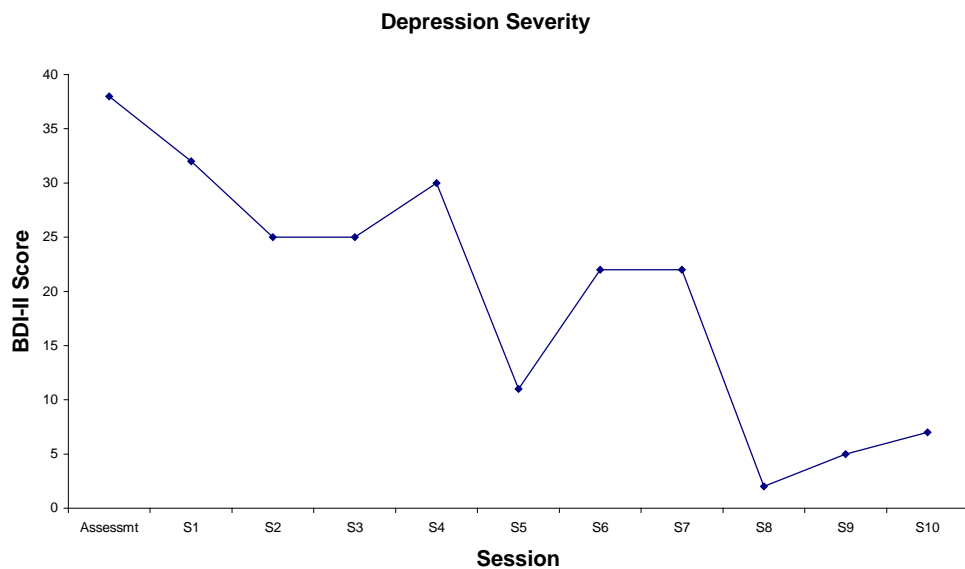
Alliance

<u>Total Alliance</u>

The Total Alliance score gradually decreased 10 points on the WAI-SR-O in a stepwise fashion over the first two weeks of therapy. It fluctuated over S 5 and S 6 and then remained fairly stable to S 8, before gradually decreasing again in S 9 and S 10. Throughout the 10 sessions the Total Alliance score stayed in the high range of the WAI-SR-O with the scores ranging between 40 and 50.5.

Graph showing Total Alliance Scores for Karen
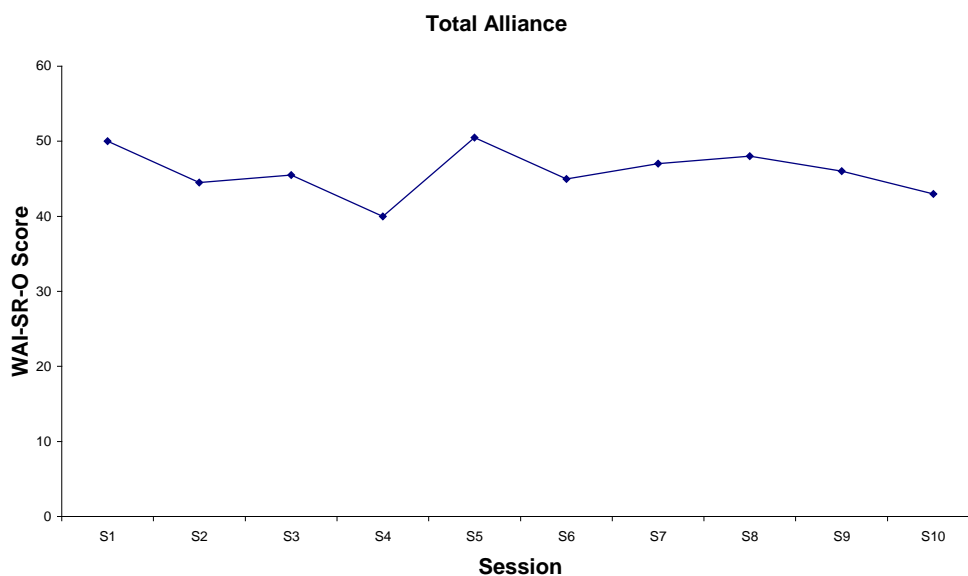
<u>Goal Subscale</u>

The Goal subscale scores fluctuated each session, gradually decreasing until S 4, before sharply increasing in S 5 and then remaining fairly stable for the remainder of the ten sessions.

<u>Task Subscale</u>

Task subscale scores followed a similar fluctuating pattern as the Goal subscale scores until S 7. From S 8 to S 10 it appeared that if the Goal subscale score increased the Task Subscale score decreased and visa versa.

Bond subscale scores gradually decreased in a linear fashion over the first two weeks of therapy. They increased in S 5 and then sharply dropped to a score of 12 on the WAI-SR-O subscale in S 6. The Bond subscale scores gradually increase again until S 8 before gradually decreasing again to S 10. The Bond dimension of the alliance was the most developed over the ten sessions, with the exception of S 6.

Graph displaying Goal, Task and Bond Subscale Scores for Karen.

Alliance Subscales
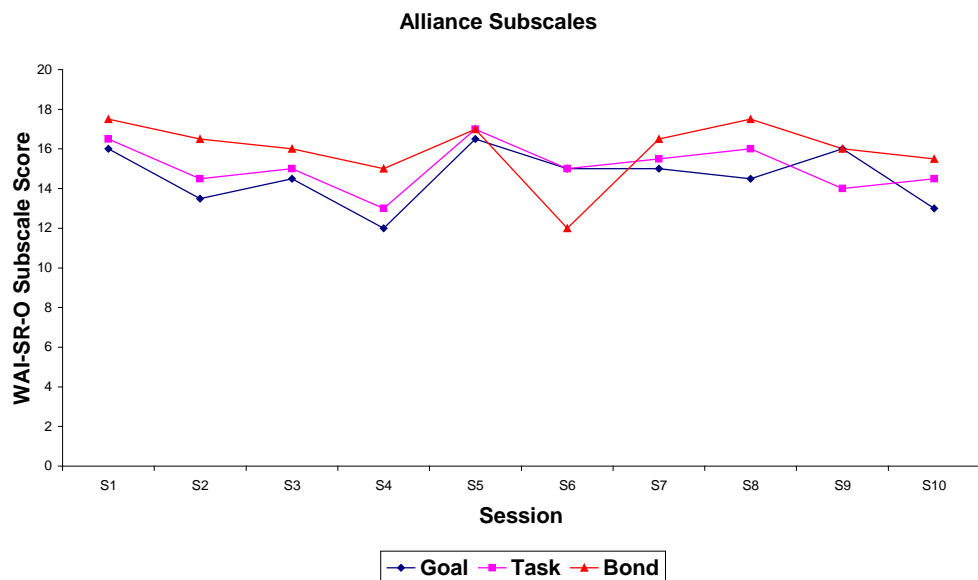
Alliance and Symptomatic Change

In the first two weeks of therapy the Total Alliance score decreased and the severity of Karen's depression decreased. However, in S 5 there was an increase in Total Alliance and all the subscales scores and a sharp decline in Karen's depression severity. In S 6 there was another increase in depression severity and a decrease in Total Alliance and Bond subscale scores. From S 6 to S 8 there was a gradual increase and maintenance of the Bond subscale scores and from S 7 to S 8 there was a decrease in Karen's depression severity. There was a decrease of the Bond subscale scores from S 8 to S 10 and an increase in Karen's depression severity from S 8 to S 10

257

Case 4 ("Rochelle")

Background information

Rochelle was a 46 year old Caucasian female. She was polite, pleasant and rapport could be gained, but she was extremely sensitive to the feeling of being controlled. Rochelle would often agree to complete an activity, but then not follow through with it. She often arrived late or rescheduled therapy sessions. Rochelle was married and had five sons; the youngest was 18 years old and the oldest 31 years old.

Her presenting symptoms included: low mood, anhedonia ("I can't be bothered dressing nicely"), tearfulness, weight loss, initial insomnia and waking up to noises very easily, indecisiveness, inner turmoil, poor memory, and she could no longer concentrate on her craft making. She felt guilty and worthless and was exhausted from having to put a face on at work and trying to portray herself as a strong person. Rochelle reported that she has been insecure all her life and was on edge all the time. Currently she was not enjoying her marriage or her life and said her husband was critical of her having "me time". He did not know that she was attending therapy. She had "no desire to die".

At initial assessment Rochelle was diagnosed with a major depressive episode and she scored 44 on the BDI-II which placed her depression severity in the severe range. The CIDI assessed Rochelle as also suffering from post traumatic stress disorder and generalized anxiety disorder.

Results

Depression Severity

At the pretreatment assessment Rochelle's depression severity was 44 points on the BDI-II. Her depression severity decreased to 18 points in S 1. The severity of Rochelle's depressive symptoms fluctuated dramatically over the first two weeks of therapy. Her depression severity increased 16 points over the first week of therapy, decreased 19 points at S 3 and then increased 8 points at S 4. The severity of Rochelle's depression then appeared to stabilize, fluctuating slightly between the scores of 19 and 21 on the BDI-II from S 5 to S 9. Her BDI-II score at S 10 was 16. Rochelle dropped out of therapy after Session 13. She went overseas and on her return she decided that she did not want to continue therapy. Her final BDI-II score was 15 (i.e., mild depression).

Graph displaying Rochelle's Depression Severity Scores
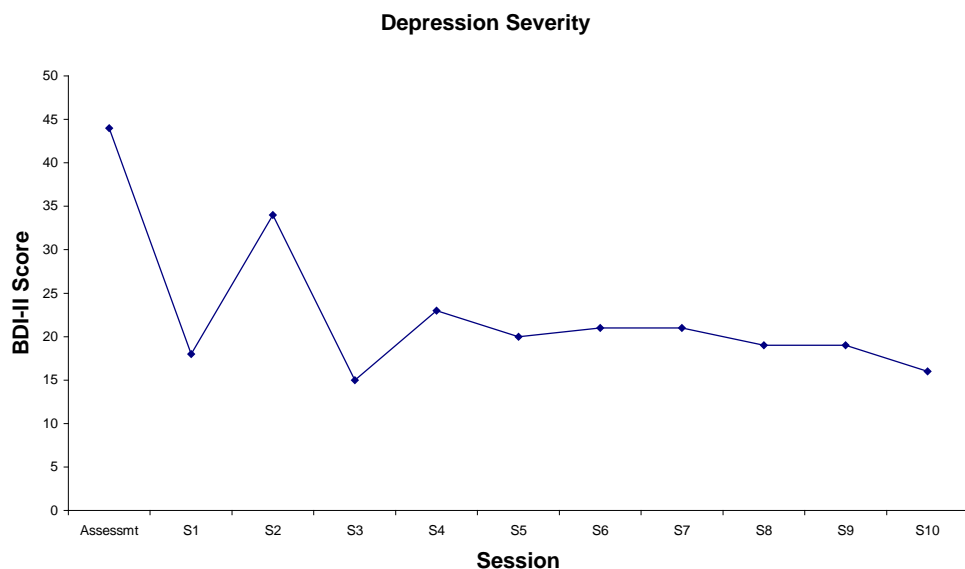
Alliance

Total Alliance

The Total Alliance score fluctuated weekly over the 10 sessions. It gradually increased from S 1 to S 3, gradually decreased to S 5, sharply increased in S 6, and then gradually decreased to S 9 before sharply increasing again in S 10.Throughout the 10 sessions the Total Alliance score stayed in the high range of the WAI-SR-O with the scores ranging between 43.5 and 51.

Graph showing Total Alliance Scores for Rochelle.
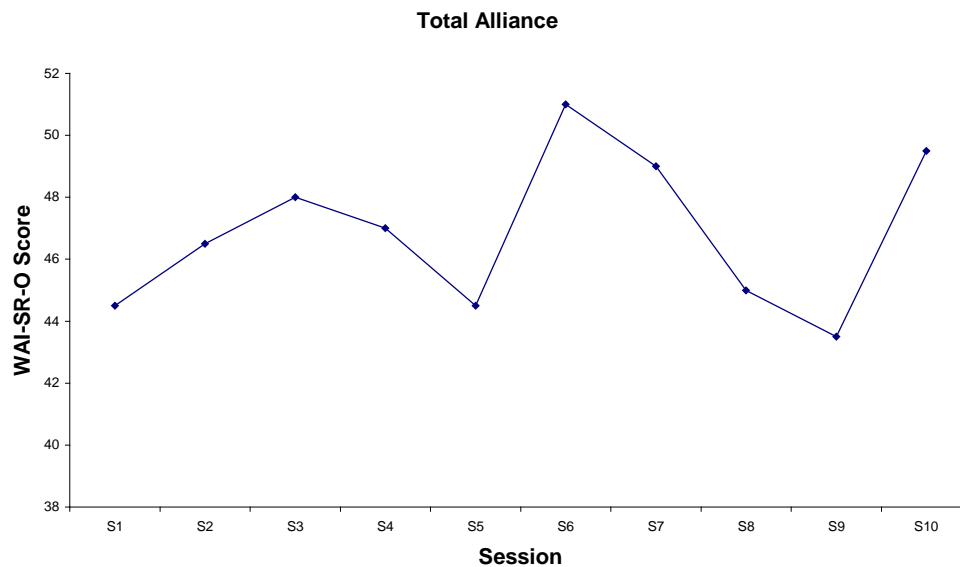
Goal Subscale

The Goal subscale score gradually increased in a fluctuating manner to S 7, decreased from S 8 to S 9 before increasing again at S 10. The Goal dimension of the alliance was the most developed between S 2 and S 7.

Task Subscale

Task subscale score decreased in the first week of therapy and then it increased gradually in a gentle fluctuating manner throughout the remainder of the sessions.

Bond Subscale

The Bond subscale score gradually increased over the first three sessions, decreased sharply at S 4 before gradually increasing again in a fluctuating manner for the remainder of the sessions.

Graph displaying Goal, Task and Bond Subscale Scores for Rochelle.

Alliance Subscales
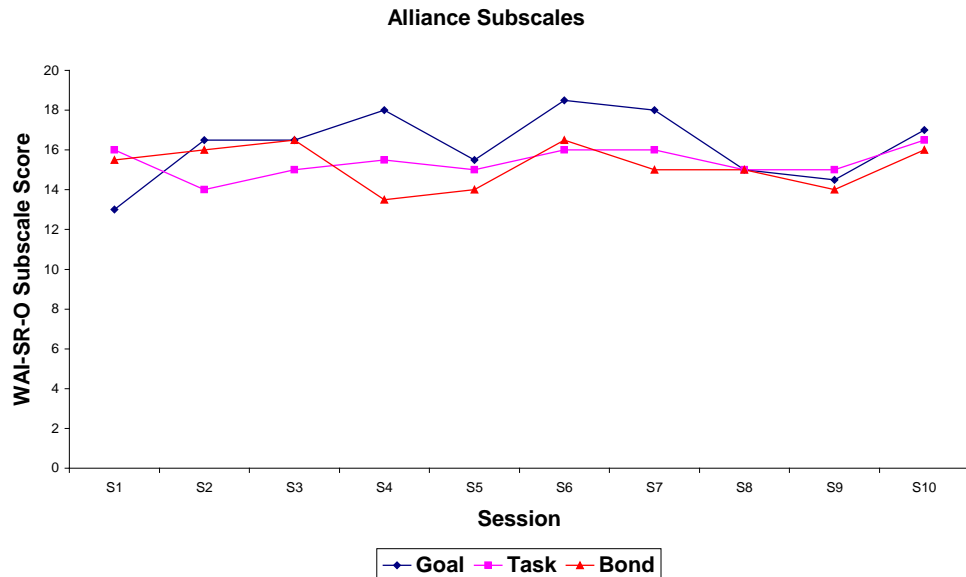
Alliance and Symptomatic Change

There was a gradual increase in the Total Alliance score over the first two sessions and a decrease in severity of depressive symptoms in the third session. There was a lowering of the Bond subscale score in S 4 and an increase in Rochelle's depression severity in that session. There did not appear to be a clear relationship between alliance and her depressive syptoms during the latter half of the rated sessions. However, there was an increase and gradual building of the Task and Bond subscale scores between S 5 and S 10 and Rochelle's depressive symptoms were more stable over these sessions.

Case 5 ("Susan")

Background information

Susan was a 37 year old Caucasian female. She was casually but tidily dressed and slightly overweight. Susan was married and has three boys, 4 years 6 months, 2 years 6 months and 1 year old. She was easily engaged in therapy.

Susan presented with symptoms of loss of interests, tearfulness, indecisiveness, and lack of concentration. She was not sleeping well due to continually being woken up by her youngest child. Once awake she had difficulty going back to sleep and lay in bed thinking about all the things she had to do. She had lost weigh recently without trying to do so. The older two children squabbled a lot and by the end of the day she was screaming at them. She then felt guilty and berated herself for not being a better parent. Toward the end of her last pregnancy she often felt frustrated at not being able to physically run after them. Her work load increased when her youngest child was born and she became overwhelmed when she tried to take all three of them out. It became increasingly more difficult for her to keep them all under control, once her youngest became more active. Finding "me time" had also been difficult. Susan and her family shifted to Auckland when she was pregnant with their youngest child so currently she was socially isolated and had no major supports. She felt unloved and unworthy. She started having suicidal thoughts 2-3 months ago, but had no plan. She said that she knew the thoughts would pass.

At the initial assessment Susan was diagnosed with major depressive disorder. She scored 19 on the BDI-II which placed her at the higher end of the mild range for depression severity.

262

Results

Depression Severity

At the pretreatment assessment Susan's depression severity was 19 points on the BDI-II. Her depression severity was 14 points in S 1. The severity of Susan's depressive symptoms fluctuated over the 10 sessions. It gradually increased until S 3, decreased 5 points in S 4 and then a further 3 points at S 5 before increasing 5 points in S 6. Her BDI- II scores then fluctuated in a more stable manner between the scores of 9 and 13 over the next four sessions. Susan completed 17 sessions with her final score on the BDI-II being 3 (i.e., no depression). At both follow up sessions Susan scored 0 on the BDI-II.

Graph displaying Susan's Depression Severity Scores

Depression Severity

BDI-II Score vs *Session*
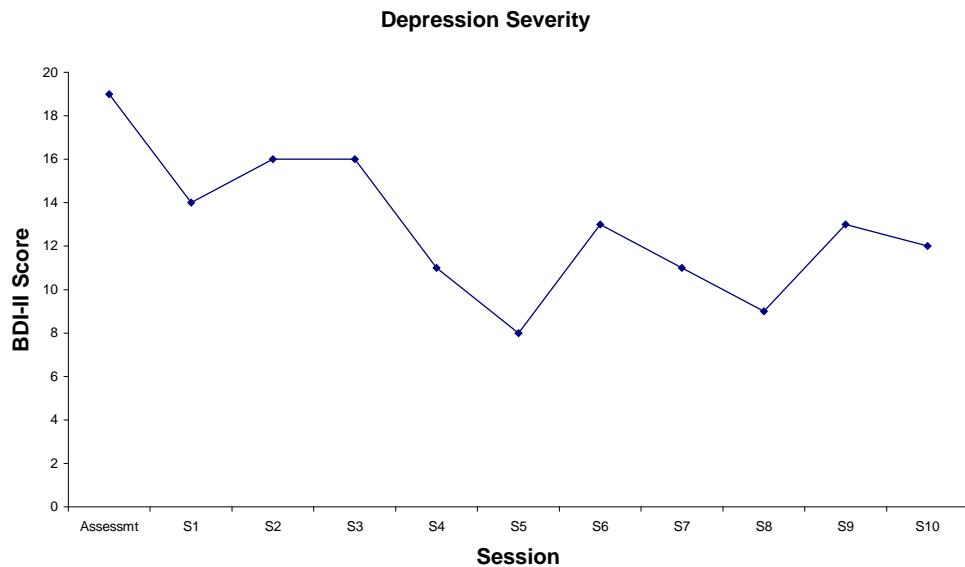
Alliance

<u>Total Alliance</u>

The Total Alliance score tended to fluctuate dramatically over the ten sessions. The Total Alliance score increased from S1 to S3, decreased sharply to S 5, before increasing again just as sharply in S 6. They continued to increase in a fluctuating manner from S 6 to S 9 before dropping sharply at S 10. Throughout the ten sessions the Total Alliance score stayed in the high range of the WAI-SR-O with the scores ranging between 41.5 and 47.

Graph showing Total Alliance Scores for Susan.

Total Alliance
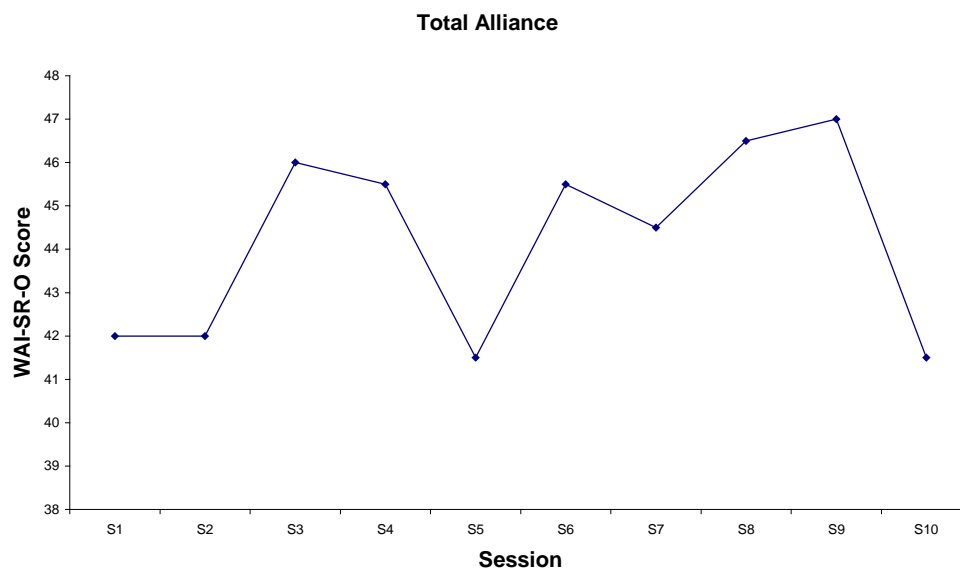
<u>Goal Subscale</u>

The Goal subscale score gradually increased in a stable pattern between S1 and S 5. It then gradually decreased to S 8, increased in S 9 and then decrease again in S 10.

<u>Task Subscale</u>

The Task subscale score gently fluctuated in a stable but gradually increasing line throughout the 10 sessions.

Over the first two weeks of therapy Bond subscale scores gradually increased. However, in S 5 there was a steep decline in the Bond subscale score. From S 6 to S 8 it gradually increased again and then gradually decreased to S 10. The Bond dimension of the alliance was the most developed throughout most of the 10 sessions

Graph displaying Goal, Task and Bond Subscale Scores for Susan.
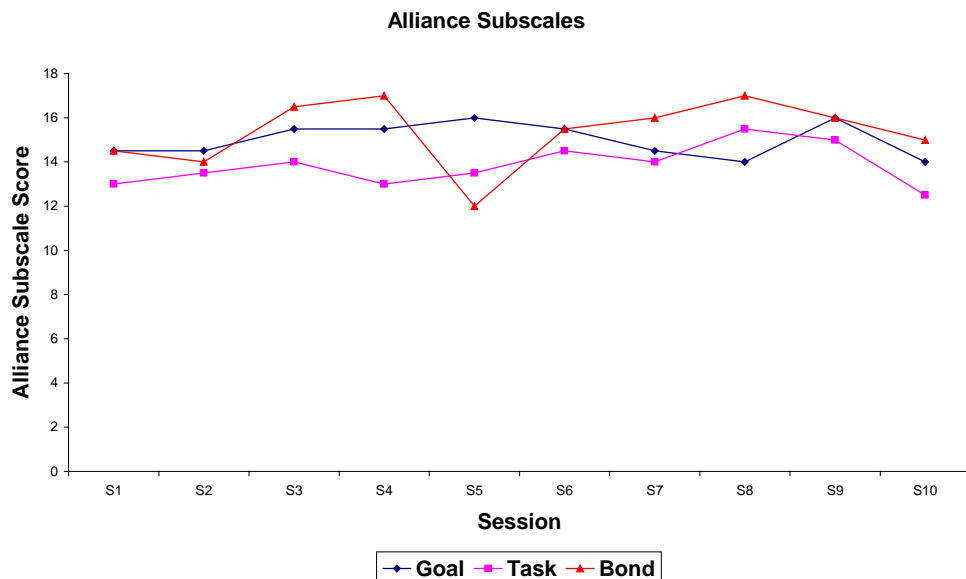
Alliance and Symptomatic Change

There was an increase of the Total Alliance and Bond subscale scores from S2 to S 3, an increase in the Goal subscale score from S 1 and S 5 and a decrease in depression severity from S 3 to S 5. There was also a gradual building of the Total Alliance and Bond subscale scores from S 6 to S 8 and a decrease in depression severity in S 7 and S 8. There was a decline in Total Alliance and Bond subscale scores in S 5 and an increase in the severity of Susan's depressive symptoms in S 6.

Case 6 ("Christine")

Background information

Christine was a 51 year old friendly Caucasian female with whom rapport was easily established. She could be jovial, witty and had a good sense of humour. She was tall, slim and well groomed. Christine was married and had three sons; 27, 26 and 22 years old. She had another unplanned baby in her forties, after which she had post natal depression. Her daughter is now 19.5 years. Therapy sessions needed to be fitted around her work commitments. Christine could easily express herself cognitively, but sometimes her affect did not match her emotions. She was responsive to praise and often sought reassurance.

Christine could not recall what precipitated her depressive symptoms which included: overwhelming sadness, her mood was lower than normal, tearfulness (worse over the last two weeks), anhedonia, initial and middle insomnia, worry, fatigue, indecisiveness (not at work, but at home), memory problems, loss of confidence, and increased appetite (she had gained weight). Christine reported that her mood was negative and that she felt unsuccessful and not good at anything. She avoided social interactions and doesn't like being around people at the moment.

At the initial assessment Christine was diagnosed with major depressive disorder. She scored 26 on the BDI-II which placed her in the moderate range for depression severity. The CIDI assessed Christine as also suffering from generalized anxiety disorder.

Results

Depressive Severity

At the pretreatment assessment Christine's depression severity was 26 points on the BDI-II. Her depression severity decreased to 30 points in S 1. Over the first week and a half of therapy Christine's depressive symptoms decreased dramatically with her BDI-II score dropping to 7 in S 3. The severity of her depressive symptoms then increased in S 4 to 19 on the BDI-II. In the third and fourth week of therapy her depression severity decreased gradually to a score of 6 in S 6, before stabilizing until S 9 where it increased to 16 points. However, this score decreased again to a score of 6 at S 10. Christine continued therapy until session 20 and her final BDI-II score was 2 (i.e., no depression).

Graph displaying Christine's Depression Severity Scores
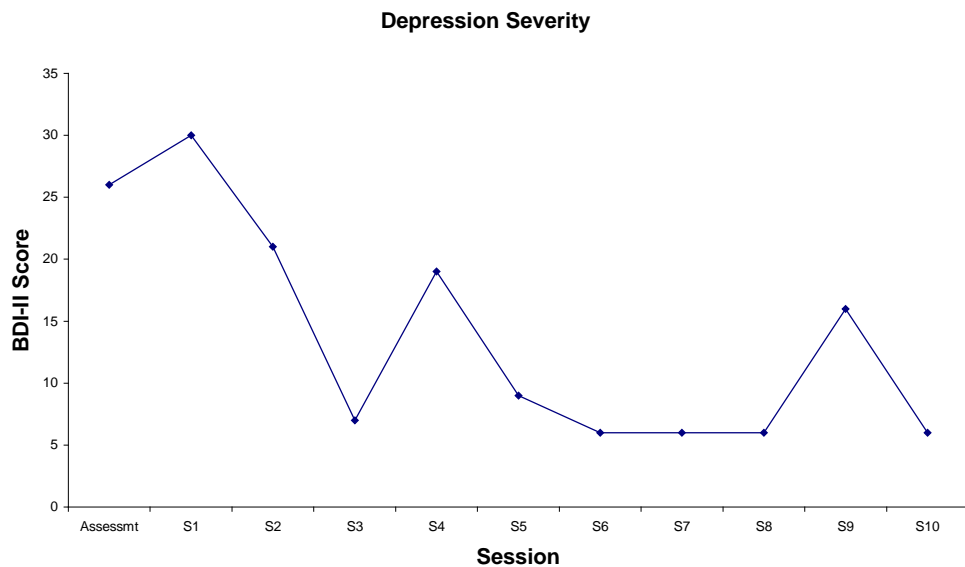
Alliance

Total Alliance

The Total Alliance score decreased over the first three sessions of therapy. It increased in S 4, and then fluctuated in a fairly stable fashion to S 10. Throughout the 10 sessions the Total Alliance score stayed in the high range of the WAI-SR-O score with scores ranging between 42 and 53.

Graph showing Total Alliance Scores for Christine.

Total Alliance
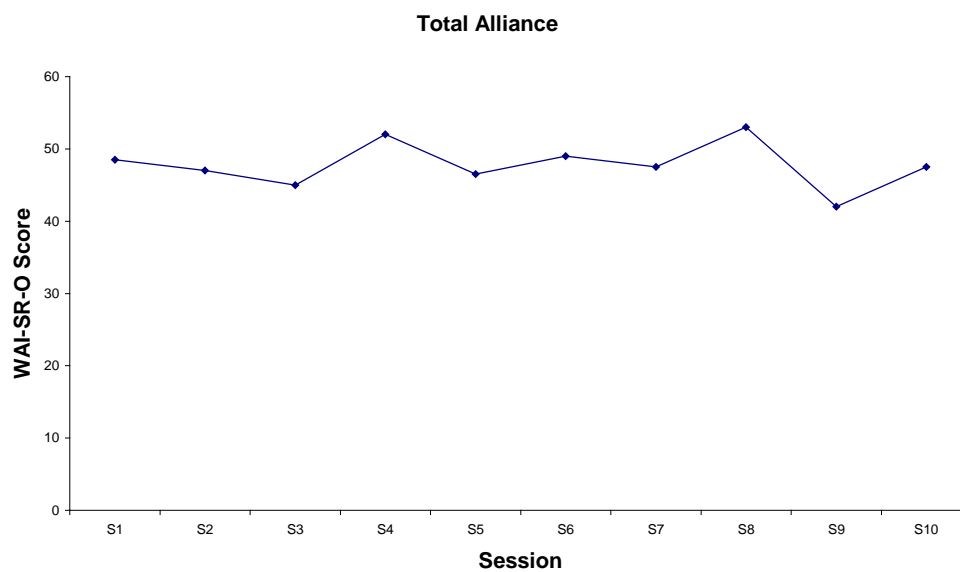
Goal Subscale

The Goal subscale score fluctuated over the first five sessions, gradually increased between S 5 and S 8, before fluctuating again to S 10.

Task Subscale

Task subscale scores followed a similar fluctuating pathway as the Goal subscale scores over the first five and last three sessions. Between S 4 and S 7 the Task subscale score gradually decreased.

Bond Subscale

The Bond subscale score fluctuated over the ten sessions. While both the Goal and Task subscale score increased in S 10 the Bond subscale score decreased. The Bond dimension of the alliance appeared to be the most developed.

Graph displaying Goal, Task and Bond Subscale Scores for Christine.
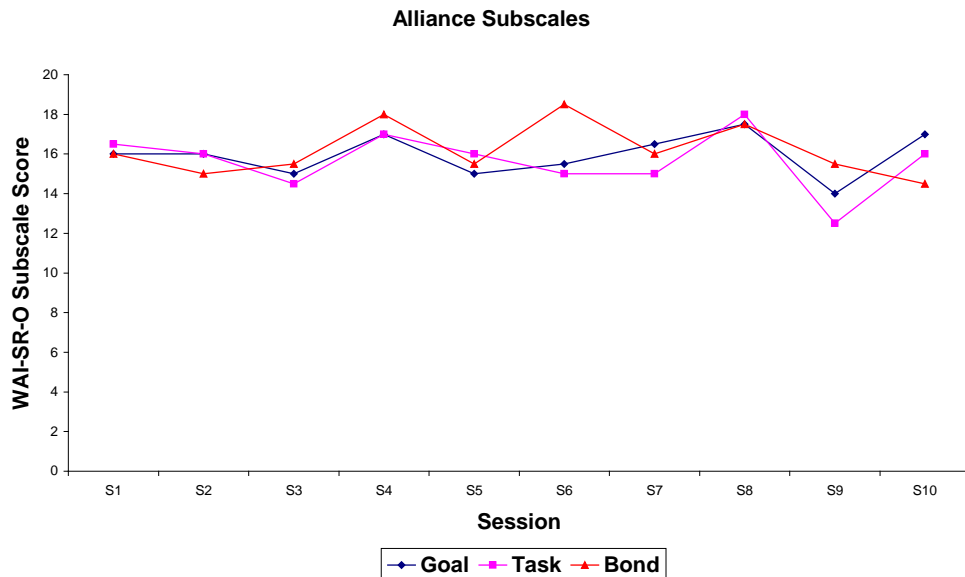
Alliance Subscales

Alliance and Symptomatic Change

There was a decrease in the Total Alliance and Goal and Task subscale scores over the first three sessions and an increase in depression severity in S 4. There was as increase and then small fluctuations in the Total Alliance and Bond subscale scores from S 5 to S 8 and a stabilization of depressive symptoms over these sessions. In S 10 there was an increase in the Total Alliance and Goal and Task subscale scores and a decrease in Christine's depressive symptoms.

Case 7 ("Michelle")

Background information

Michelle was a 31 year old married Caucasian female. She was married and had two sons, 5 and 7 years old. Michelle was well groomed, but tended to have food spills on her clothes from "looking after her little nephew. Michelle had a friendly personality, but she appeared flat and was tearful. Michelle reported that her depressive symptoms seemed to appear after her best friend shifted to the South Island to live. She was also experiencing financial stress.

Her presenting symptoms included: depressed mood, anhedonia, increased appetite (she snacks more), increased need for sleep, low energy levels, procrastination, irritability, low self esteem, lack of libido and guilt about her childrens' behaviour. Both her husband and her best friend have noticed that she has physically slowed down. She reported that she tended to worry about everything and this has become worse over the last 3 months. Over the last 2 -3 weeks she had suffered from a lot of headaches, tightness in the chest and pounding heart, particularly when she was "wound up". She did not have any suicidal thoughts, but "if something would kill me that would be nice".

At the initial assessment Michelle was diagnosed with major depression disorder. She scored 23 on the BDI-II which placed her depression severity in the moderate range. The CIDI assessed Michelle has also experiencing generalized anxiety disorder and panic disorder without agoraphobia.

Results

Depressive Severity

At the pretreatment assessment Michelle's depression severity was 23 points on the BDI-II. Her depression severity decreased to 36 points in S 1. The severity of Michelle's depressive symptoms appeared to fluctuate each week of therapy. Over the first week and a half of therapy Michelle's depressive symptoms decreased to a score of 29 in S 3. The severity of her depressive symptoms then increased 11 points in S 4 to a score 40. In the third week of therapy her depression decreased 8 points to a score of 32 in S 5 and then to a score of 29 in S 6. However, this score increased again over the fourth week to 38 in S 8. In S 9 Michelle's depressive symptoms decreased to a score of 24 on the BDI-II and in S 10 her BDI-II score was 23. Michelle completed 20 sessions of therapy and her final BDI-II score was 11 (i.e., no depression).

Graph displaying Michelle's Depression Severity Scores
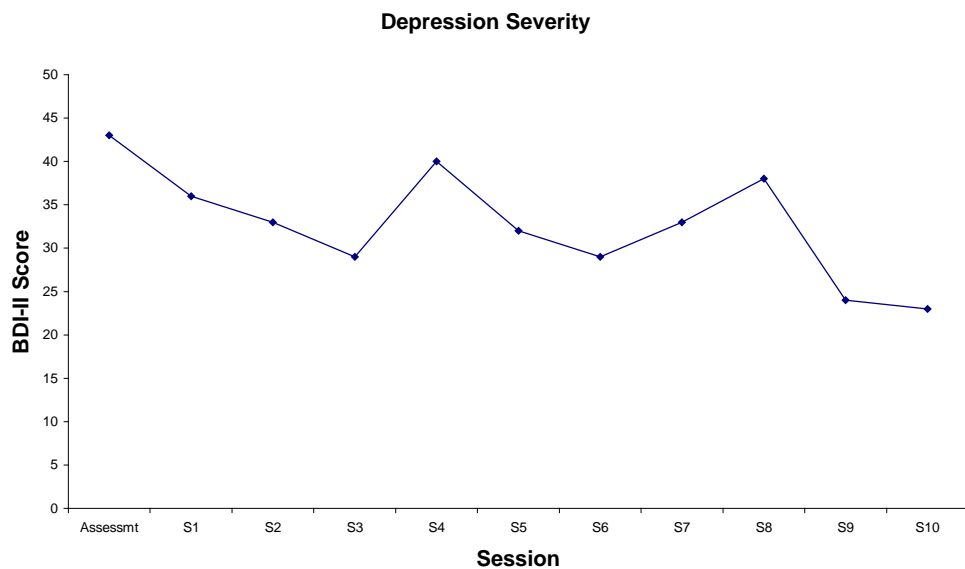
Alliance

<u>Total Alliance</u>

The Total Alliance score fluctuated in a stable line each session throughout the 10 sessions of therapy. The Total Alliance score increased from 40 .5 in S 1 to 48.5 in S 2 on the WAI-SR-O. This score then fluctuated between 45 and 52 over the next seven sessions. In S 10 the Total alliance score dropped 5.5 points to 39.5. Throughout the ten sessions the Total Alliance score stayed in the high range of the WAI-SR-O with the scores ranging between 39.5 and 49.5.

Graph showing Total Alliance Scores for Michelle.

Total Alliance
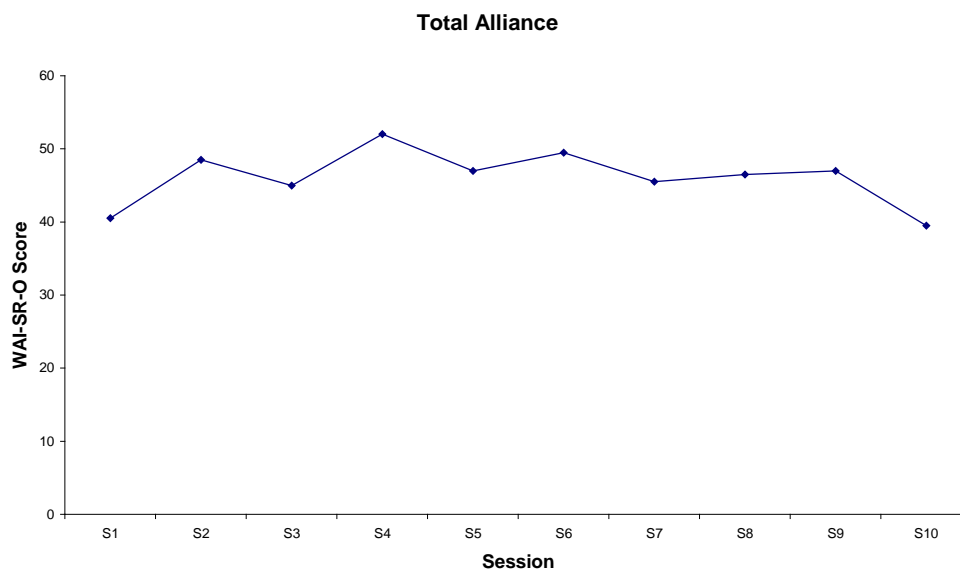
<u>Goal Subscale</u>

The Goal subscale score increased quite substantially from 10.5 in S 1 to 16 in S 2. It then fluctuated each session until S 6, gradually decreased to S 8, increased slightly in S 9 before decreasing again at S10.

<u>Task Subscale</u>

The Task subscale score was mostly stable until S 8. It increased to S 9, peaking at a score of 17.5 and then decreased again at S 10 to a score of 13.5. The Task dimension of the alliance was the most developed in S 9.

The Bond subscale score gradually increased over the first two weeks of therapy before gradually decreasing to S 10. The Bond dimension of the alliance was the most developed throughout the first eight sessions.

Graph displaying Goal, Task and Bond Subscale Scores for Michelle.
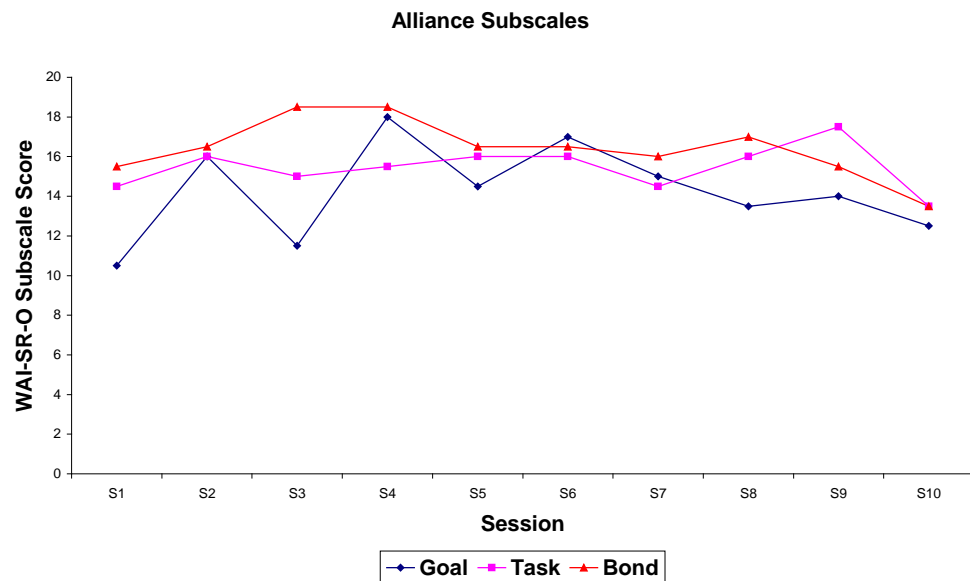
Alliance and Symptomatic Change

There was a decrease in Michelle's depression severity in S 2 and S 3, S 5 and S 6 and S 9 and an increase in the Total Alliance scores in S 2, S 4, S 6 and S 9. There was an increase in Michelle's depression severity in S 4, S 7 and S 8 and a decrease in Total Alliance scores in S 3, S 5, S 7 and S10. The Bond subscale scores gradually increased from S 1 to S 4 and Michelle's depression severity gradually decreased from S 1 to S 3. The Task subscale score increased in S 8 and S 9 and there was a decrease in Michelle's depression severity in S 9 and S 10.

Case 8 ("Sarah")

Background information

Sarah was a 47 year old Caucasian female. She was married with two children who were 12 and 16 years old. Sarah was well groomed most of the time. Rapport was good and she maintained eye contact, but her mood was labile.

Her presenting symptoms included: anhedonia, indecisiveness, loss of libido, restlessness, increased appetite, lack of concentration and exhaustion. Sarah woke up early in the morning worried and anxious. She ruminated over choices and felt more anxious when she was under pressure to make a decision. She perceived that no matter what she decided it would be the wrong decision. She also worried about being a terrible mother and that she was a failure. She had difficulty coping with everyday tasks, stresses over mess and then becomes overwhelmed because she hasn't done anything. She could feel guilty because nothing catastrophic was happening ("I can't help myself"). She perceived that her outlook was bleak. This year she had noticed that she has had trouble throwing things out because "she may need it" and she always has to check around the car to make sure she hasn't dropped anything. She sometimes had thoughts that her family would be better off without her. Sarah had a part time job in retail and she liked being out and about with other people. She has noticed that she was less depressed when she was away from the house. At work she can put on a mask, whereas she could not in her personal life.

At the initial assessment Sarah was diagnosed as having a major depressive episode. She scored 53 on the BDI-II which placed her depression severity in the severe range. The CIDI assessed Sarah as also suffering from obsessive-compulsion disorder, post traumatic stress disorder and generalized anxiety disorder. Sarah's depressive symptoms appeared to be the most prominent at the initial assessment, but further into therapy it was realized that she had strong OCD and GAD traits which needed therapeutic attention.

Results

Depression Severity

At the pretreatment assessment Sarah's depression severity was 53 points on the BDI-II. The severity of Sarah's depressive symptoms fluctuated each week of therapy. It decreased from 52 in S 1 to 48 in S 2. In the second week of therapy the severity of her depressive symptoms increased 16 points to 54 on the BDI-II in S 3 and in S 4 this score decreased slightly to 53. In the third week of therapy her depression severity increased slightly to a score 54 on the BDI-II and then it dropped sharply to 48 again in S 6. The severity of Sarah's depressive symptoms fluctuated between the scores of 46 and 50 from S 7 to S 9. In S 10 this score rose to 53. After 20 sessions of therapy, Sarah's final score of 31 still fell into the severe range of depression severity. At the 2 monthly follow-up she also scored 31 on the BDI-II. She was offered further therapy sessions through The Psychology Centre.

Graph displaying Sarah's Depression Severity Scores
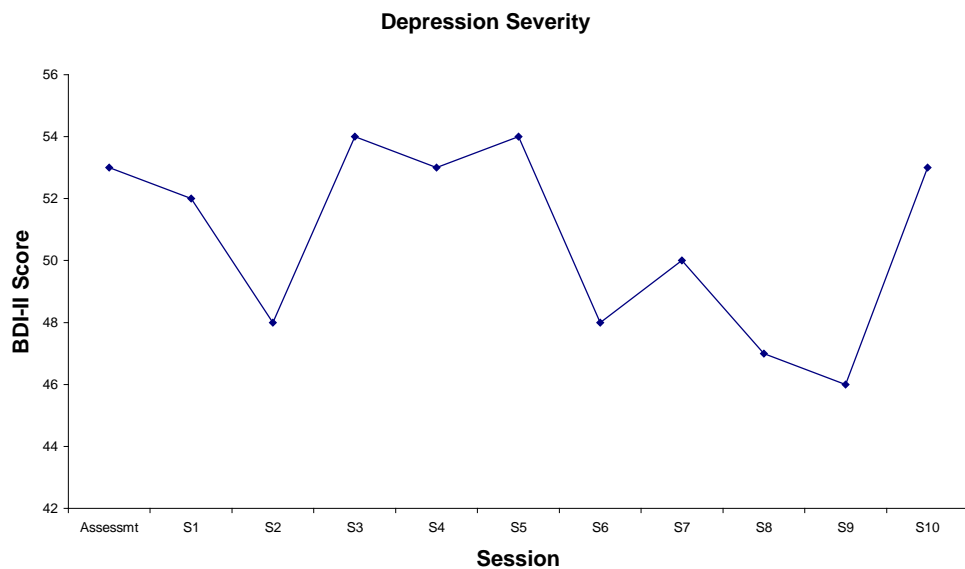
275

Alliance

<u>Total Alliance</u>
The Total Alliance score fluctuated throughout the ten sessions. It increased from a score of 41 in S 1 to 46.5 in S 2 on the WAI-SR-O. In the second week the Total Alliance score dropped 11 points to 35.5 in S 3. From S 4 they gradually increased to peak at 46 in S 6. In the fourth week of therapy and in S 9 they fluctuated between 43 and 45. The Total Alliance score decreased 10 points to 33.5 in S 10. Throughout the 10 sessions the Total Alliance scores ranged from the high end of the moderate range to the lower end of the high range of the WAI-SR-O score with the scores ranging between 33.5 and 46.5.

Graph showing Total Alliance Scores for Sarah.
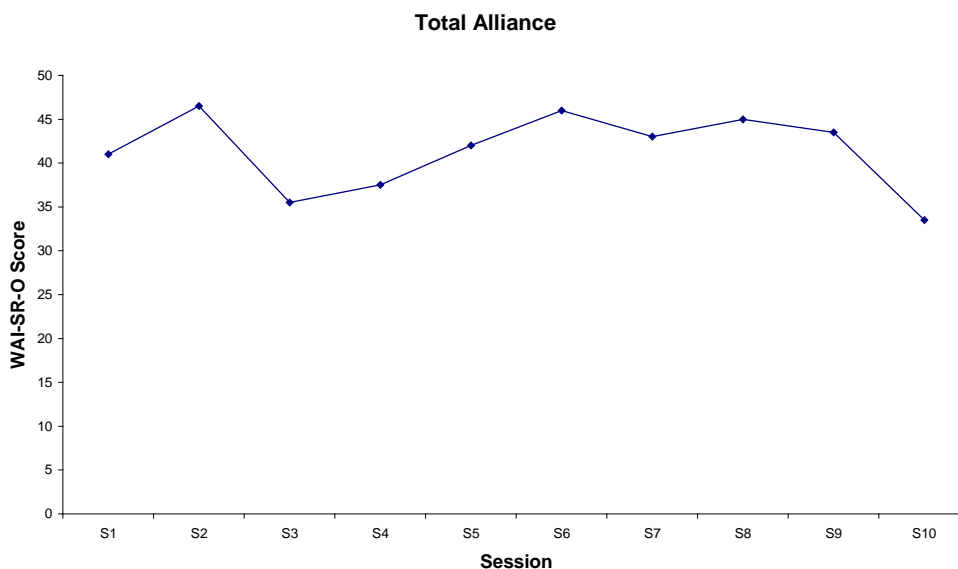
<u>Goal Subscale</u>
The Goal subscale score fluctuated throughout the ten sessions. It increased in S 2 to a score of 15.5 and then dropped sharply to a score of 7.5 in S 3. The Goal subscale score gradually increased to a score of 16 in S 6, and then dropped again to 10 at S 7, before increasing to 15 in S 9. All subscale scores decreased at S 10.

<u>Task Subscale</u>
The Task subscale appeared to follow the same fluctuating pattern of as the Total Alliance scale.

The Bond subscale score gradually decreased from 18.5 in S 1 to 15 in S 4, then gradually increased to a score of 17.5 in S 7, before gradually decreasing again to 12.5 in S10. The Bond dimension of the alliance was the most developed throughout the ten sessions.

Graph displaying Goal, Task and Bond Subscale Scores for Sarah.

Alliance Subscales
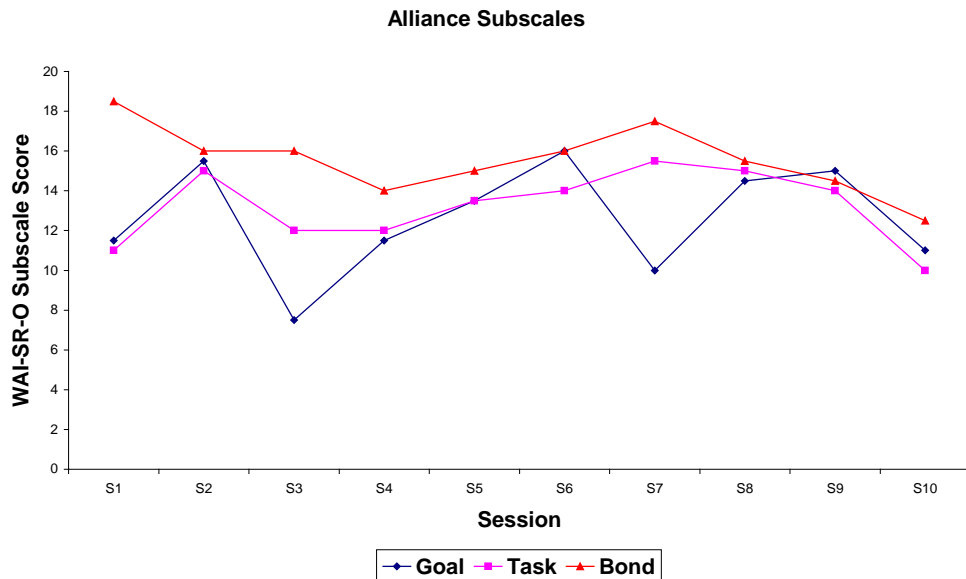
Alliance and Symptomatic Change

There was an increase in Total Alliance scores in S 2 and a gradual increase of Total alliance scores from S 3 to S 6 and a decrease in Sarah's depressive severity in S 2 and S 6. There was a decrease in the Total Alliance and Goal subscale scores in S 3 and S 7 and an increase in Sarah's depressive severity from S 3 to S 5.

Case 9 ("Dennis")

Background information

Dennis was a 57 year old Caucasian male. He was tall, well groomed, but causally dressed. Rapport was good and he maintained good eye contact. Dennis was employed as a technician for 30 years and then worked for a company that repaired photocopiers, before retiring. Currently he was in a defacto relationship and had two adult children.

His presenting symptoms included: sadness, decreased pleasure, poor concentration, memory problems (i.e., can't remember people names and dates), middle insomnia, weight loss, lack of motivation and indecisiveness. Dennis worried about "everything" and had a fear that "someone, somewhere was coming for him". He was hypervilgilant and easily startled. He tended to get angry, particularly when he was driving. He reported that he had no suicidal thoughts, but "it wouldn't worry him if he dropped dead tomorrow". He used to drink 2-3 glasses of wine per week, whereas currently he and his partner were drinking 3 glasses of wine per night. However, Dennis did not report any alcohol dependency or abuse.

At the initial assessment Dennis was diagnosed with major depression symptomology. He scored 30 on the BDI-II which placed him at the lower end of the severe range for depression severity. The CIDI also assessed Dennis as also currently experiencing post traumatic stress disorder, generalized anxiety disorder and panic disorder without agoraphobia.

Results

Depression Severity

At the pretreatment assessment Dennis' depression severity was 30 points on the BDI-II. The severity of Dennis's depressive symptoms increased over the first week of therapy from 28 in S 1 to 30 in S 2. In the second week of therapy there was a sharp drop of 11 points in his depression severity and then a slight increase to a score of 22 on the BDI-II in S 4. In the third week of therapy his depression severity progressively decreased from a score 20 in S 5 to 16 in S 6. The severity of Dennis' depressive symptoms stabilized at this score until S 9. It decreased to a score of 8 at S 10. Dennis completed 20 sessions of therapy and his final BDI-II score was 4 (i.e., no depression).

Graph displaying Dennis' Depression Severity Scores
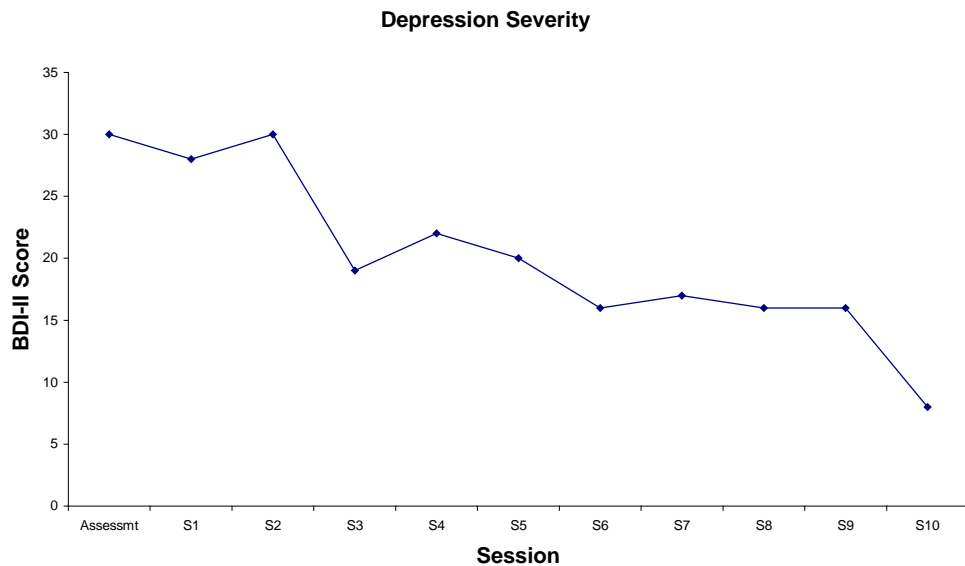
Alliance

Total Alliance

The Total Alliance fluctuated in a stable line between 40.5 and 47.5 on the WAI-SR-O over the first nine sessions. In S 10 it dropped 7 points from a score of 43.5 to 36.5. Throughout the 10 sessions the Total Alliance scores mostly fell into the high range of the WAI-SR-O score with the scores ranging between 40.5and 47.5. Only the last score of 36.5 fell into the top end of the moderate range.

Graph showing Total Alliance Scores for Dennis.

Total Alliance
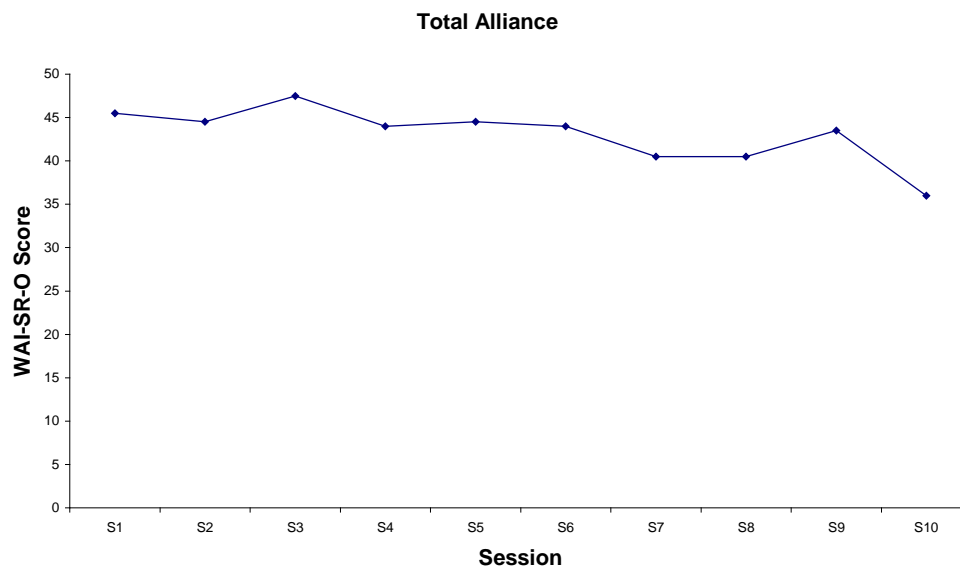
Goal Subscale

The Goal subscale score increased from a score of 14.4 in S 1 to a score of 16 in S 2 and then it gradually declined in a step-wise manner to a score of 11.5 in S 8. In S 9 it increased to a score of 15 and then decreased again to a score of 11 in S 10.

Task Subscale

The Task subscale score decreased in the first week of therapy and then gradually increased in a fluctuated manner throughout the remainder sessions. From S 5 onwards the Task dimension of the alliance became the most developed. Task subscale scores gradually increased from a score of 15 in S 5, peaking at a score of 16 in S 8, before gradually decreasing to 15.5 in S 10.

The Bond subscale score fluctuated over the first three sessions. It then gradually decreased, initially in a stepwise and then liner manner, from S 3 to a score of 12 at S 10. The Bond dimension of the alliance was the most developed subscale from S 1 to S 5 with the exception of S 2.

Graph displaying Goal, Task and Bond Subscale Scores for Dennis.

Alliance Subscales
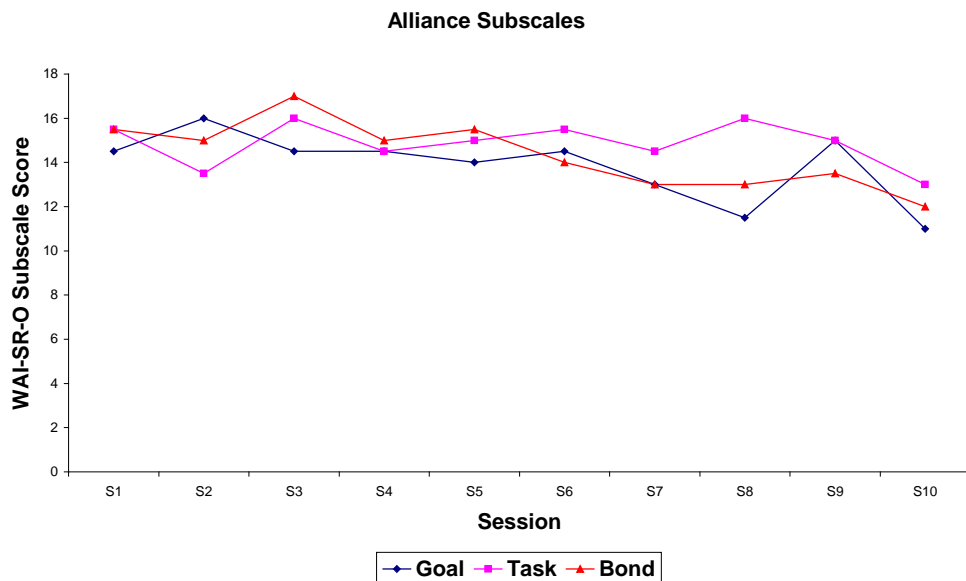
Alliance and Symptomatic Change

There was an increase in the Total Alliance and Bond and Task subscale scores from S 2 to S 3 and a decrease in Dennis' depression severity in S 3. There was decrease in these alliance scores in S 4 and an increase in Dennis' depression severity in S 4. However, after S 3 scores of the Bond dimension gradually decreased and this did not appear to influence depression severity. Scores of the Task subscale gradually increased from S 2 to S 8 and the severity of Dennis' depressive symptoms gradually declined from S 4 to S 6, stabilized from S 7 to S 9 and then sharply decreased in S 10. The Goal subscale scores increased from S 8 to S 9 and there was a decrease in his depression severity from S 9 to S 10.

Case 10 ("Rosemary")

Background information

Rosemary was a 49 year old Caucasian solo mother with four children. She was divorced from her youngest daughter's father. Rosemary didn't take a lot of care with her appearance, but she was clean and tidy. It was harder to gain rapport with Rosemary and eye contact was only maintained sometimes. She was deeply embarrassed abut her current situation.

Her presenting symptoms included: low mood, lack of motivation and enjoyment, decreased appetite, insomnia (intermittent waking), blank mind, forgetfulness, and difficulty making decisions. She felt physically slow, had low self esteem, found it difficult to cope and believed that she had no future. She was pessimistic and thought that nothing was going to work and that she was stuck in a horrible situation. However, she was not suicidal.

At the initial assessment Rosemary was diagnosed with major depression disorder. She scored 32 on the BDI-II which placed her depression severity at the lower end of the severe range. The CIDI assessed Rosemary as also experiencing generalized anxiety disorder.

Rosemary terminated therapy after Session 6.

Results

Depression Severity

At the pretreatment assessment Rosemary's depression severity was 32 points on the BDI-II. In the first week of therapy the severity of Rosemary's depression decreased from 22 in S 1 to 19 in S 2. It week two it increased 2 points to 21, then decreased gradually throughout the remainder of her sessions. Rosemary's final BDI-I score was 14 at S 6 (i.e., mild depression).

Graph displaying Rosemary's Depression Severity Scores
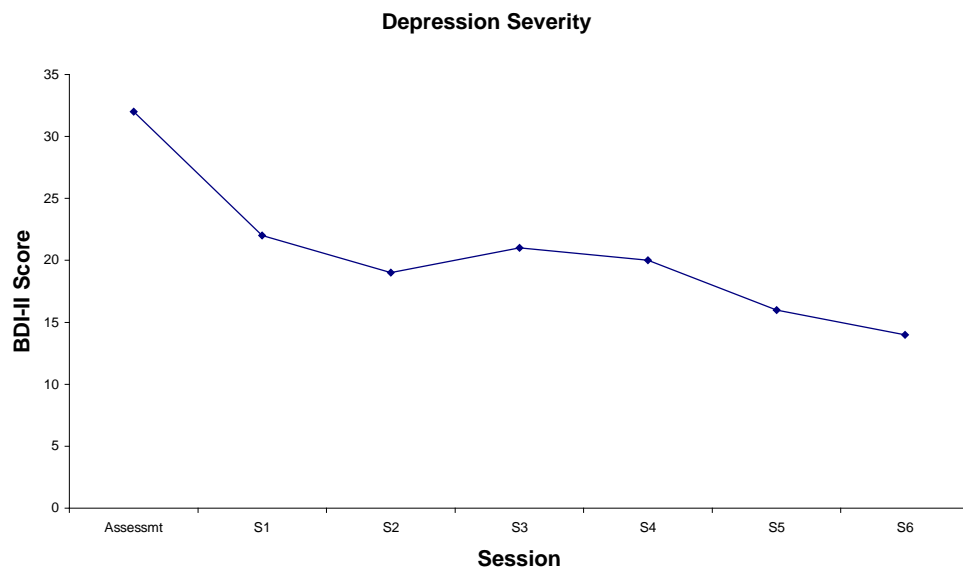
Alliance

Total Alliance

The Total Alliance score appeared to fluctuate in a stable line throughout the six sessions. The Total Alliance scores increased in the first week from 38 in S 1 to 43 in S 2. It then decreased in S 3, increased again to S 5, before decreasing again in S 6. Throughout the 6 therapy sessions the Total Alliance scores fluctuated between the high end of the moderate range to the low end of the high range with the scores ranging between 35.5 and 45 on the WAI-SR-O.

Graph showing Total Alliance Scores for Rosemary.

Total Alliance
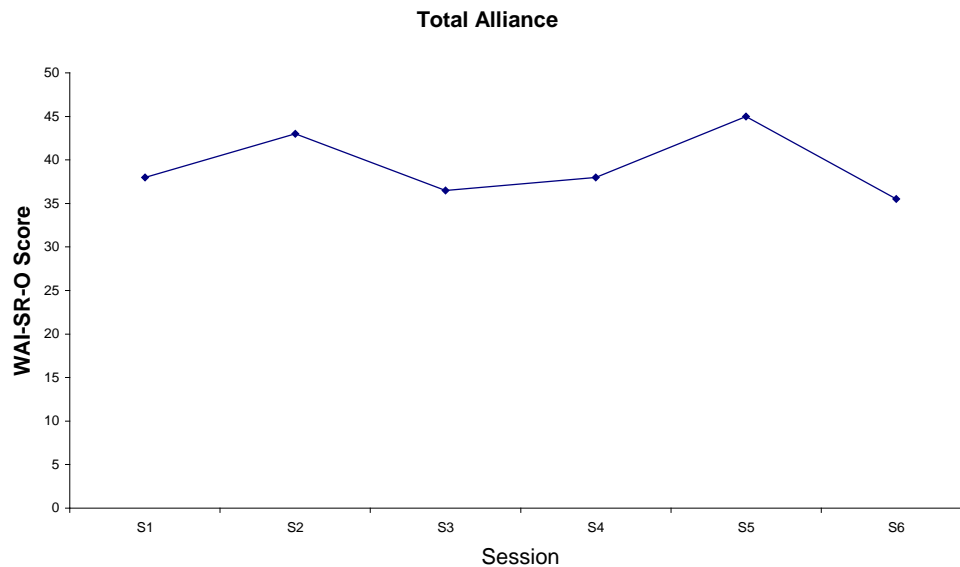
Goal Subscale

Goal subscale scores fluctuated over the first three sessions, increased to S 5 and then decreased in S 6. The Goal dimension of the alliance was the most developed in S 5.

Task Subscale

Task subscale scores increased in the first week, decreased over the second week, increased again at S 5 before decreasing in S 6. The Task dimension of the alliance was the most developed over the first four sessions.

Bond Subscale

The Bond subscale score fluctuated throughout the first three sessions, but then gradually increased to S 5. It then decreased in S 6.

Graph displaying Goal, Task and Bond Subscale Scores or Rosemary.

Alliance Subscales
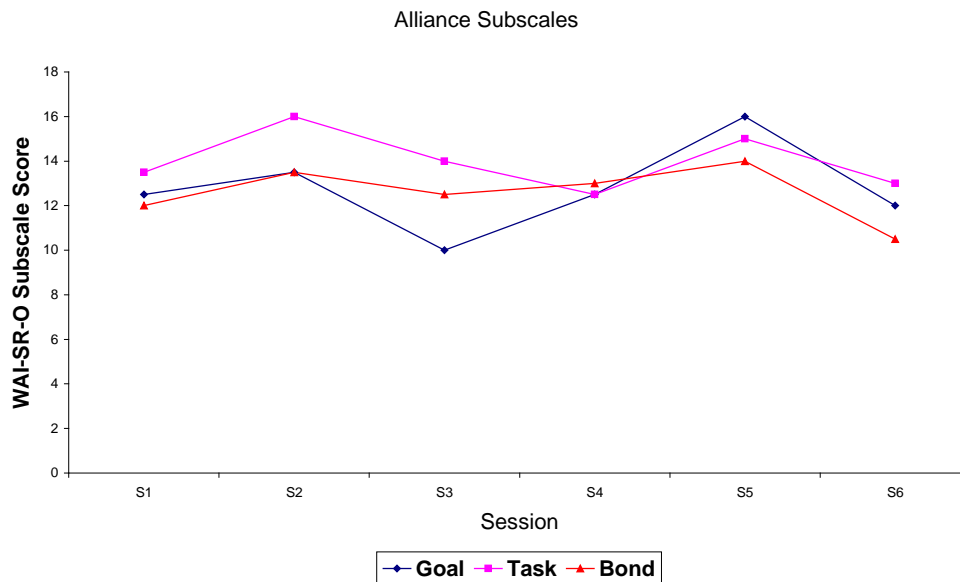
Alliance and Symptomatic Change

In the second session of therapy the severity of Rosemary's depression decreased and there was an increase in Total Alliance and Task subscale scores. From S 3 to S 5 there was an increase in the Total Alliance and Goal and Bond subscale scores and from S 4 an increase in the Task subscale score and there was a decrease in Rosemary's depression severity over these sessions. In S 6 all alliance scores and depression severity decreased.

Appendix G

Tables of Total Alliance and Subscale Scores

Total Alliance Scores of the WAI-SR-O

Total Alliance

Session	Case 1	Case 2	Case 3	Case 4	Case 5	Case 6	Case 7	Case 8	Case 9	Case 10
1	48	43.5	**50**	44.5	42	**48.5**	40.5	41	45.5	38
2	**53**	47	44.5	46.5	42	47	**48.5**	46.5	44.5	43
3	41	**50.5**	45.5	48	46	45	45	**35.5**	47.5	36.5
4	46.5	47	40	47	45.5	**52**	**52**	37.5	44	38
5	46.5	42	**50.5**	44.5	41.5	46.5	47	42	44.5	45
6	47.5	43.5	45	**51**	45.5	**49**	**49.5**	46	44	**35.5**
7	43.5	**49.5**	47	**49**	44.5	47.5	45.5	43	40.5	
8	44	45	48	45	46.5	53	46.5	45	40.5	
9	46	46	46	43.5	47	42	47	43.5	43.5	
10	41	48	43	**49.5**	41.5	47.5	39.5	**33.5**	**36**	

Subscale Scores of the WAI-SR-O

Goal

Session	Case 1	Case 2	Case 3	Case 4	Case 5	Case 6	Case 7	Case 8	Case 9	Case 10
1	16	14.5	16	13	14.5	16	**10.5**	**11.5**	14.5	12.5
2	**18**	14	13.5	**16.5**	14.5	16	16	15.5	16	13.5
3	13.5	16	14.5	**16.5**	15.5	15	**11.5**	**7.5**	14.5	**10**
4	15.5	15	**12**	18	15.5	**17**	**18**	**11.5**	14.5	12.5
5	16	13.5	**16.5**	15.5	16	15	14.5	13.5	14	16
6	16	14.5	15	**18.5**	15.5	15.5	**17**	16	14.5	**12**
7	13	13.5	15	**18**	14.5	**16.5**	15	**10**	13	
8	14.5	13	14.5	15	14	**17.5**	13.5	14.5	**11.5**	
9	15.5	13	16	14.5	16	14	14	15	15	
10	13.5	14.5	13	**17**	14	**17**	12.5	**11**	**11**	

Task

Session	Case 1	Case 2	Case 3	Case 4	Case 5	Case 6	Case 7	Case 8	Case 9	Case 10
1	13.5	13	**16.5**	16	13	**16.5**	14.5	**11**	15.5	13.5
2	**16.5**	14	14.5	14	13.5	16	16	15	13.5	16
3	13.5	16	15	15	14	14.5	15	**12**	16	14
4	16	14.5	13	15.5	13	**17**	15.5	**12**	14.5	12.5
5	15.5	12.5	**17**	15	13.5	16	16	13.5	15	15
6	14.5	12.5	15	16	14.5	15	16	14	15.5	13
7	14.5	**16.5**	15.5	16	14	15	14.5	15.5	14.5	
8	13.5	16	16	15	15.5	**18**	16	15	16	
9	14	15	14	15	15	12.5	**17.5**	14	15	
10	12.5	16	14.5	**16.5**	12.5	16	13.5	**10**	13	

Bond

Session	Case 1	Case 2	Case 3	Case 4	Case 5	Case 6	Case 7	Case 8	Case 9	Case 10
1	**18.5**	16	**17.5**	15.5	14.5	16	15.5	**18.5**	15.5	**12**
2	**18.5**	**19**	**16.5**	16	14	15	**16.5**	16	15	13.5
3	14	**18.5**	16	**16.5**	**16.5**	15.5	**18.5**	16	**17**	12.5
4	15	**17.5**	15	13.5	17	**18**	**18.5**	14	15	13
5	15	16	**17**	14	**12**	15.5	**16.5**	15	15.5	14
6	**17**	**16.5**	**12**	**16.5**	15.5	**18.5**	**16.5**	16	14	**10.5**
7	16	**19.5**	**16.5**	15	16	16	16	**17.5**	13	
8	16	16	**17.5**	15	**17**	**17.5**	**17**	15.5	13	
9	**16.5**	**18**	16	14	16	15.5	15.5	14.5	13.5	
10	15	**17.5**	15.5	16	15	14.5	13.5	12.5	**12**	

Total Alliance Scores

Scores on the WAI –SR-O

12 – 24 - Low Range (Seldom/Sometimes)

25 -36 – Moderate Range (Fairly Often)

36.5 – 48 - High Range (Very Often)

48.5 – 60 – Very High Range (Always)

Subscales Scores

Scores on WAI-SR-O

1 -4   - Very Low Range (Seldom)

4.5 – 8 – Low Range (Sometimes)

8.5 – 12 - Moderate Range (Fairly Often)

12.5 – 16 - High Range (Very Often)

16.5 – 20 – Very High Range (Always)

Depression Severity, Total Alliance and Subscales Scores of Cases.

Case 1

BDI-II Score	Total Alliance Score	Goal Subscale Score	Task Subscale Score	Bond Subscale score
28	48	16	13.5	18.5
21	53	18	16.5	18.5
24	41	13.5	13.5	14
25	46.5	15.5	16	15
24	46.5	16	15.5	15
27	47.5	16	14.5	17
25	43.5	13	14.5	16
17	44	14.5	13.5	16
17	46	15.5	14	16.5
18	41	13.5	12.5	15

Case 2

BDI-II Score	Total Alliance Score	Goal Subscale Score	Task Subscale Score	Bond Subscale score
12	43.5	14.5	13	16
11	47	14	14	19
8	50.5	16	16	18.5
10	47	15	14.5	17.5
10	42	13.5	12.5	16
9	43.5	14.5	12.5	16.5
12	49.5	13.5	16.5	19.5
10	45	13	16	16
12	46	13	15	18
9	48	14.5	16	17.5

Case 3

BDI-II Score	Total Alliance Score	Goal Subscale Score	Task Subscale Score	Bond Subscale score
32	50	16	16.5	17.5
26	44.5	13.5	14.5	16.5
25	45.5	14.5	15	16
30	40	12	13	15
11	50.5	16.5	17	17
22	45	15	15	12
22	47	15	15.5	16.5
2	48	14.5	16	17.5
5	46	16	14	16
7	43	13	14.5	15.5

Case 4

BDI-II Score	Total Alliance Score	Goal Subscale Score	Task Subscale Score	Bond Subscale score
18	44.5	13	16	15.5
34	46.5	16.5	14	16
15	48	16.5	15	16.5
23	47	18	15.5	13.5
20	44.5	15.5	15	14
21	51	18.5	16	16.5
21	49	18	16	15
19	45	15	15	15
19	43.5	14.5	15	14
19	49.5	17	16.5	16

Case 5

BDI-II Score	Total Alliance Score	Goal Subscale Score	Task Subscale Score	Bond Subscale score
14	42	14.5	13	14.5
16	42	14.5	13.5	14
16	46	15.5	14	16.5
11	45.5	15.5	13	17
8	41.5	16	13.5	12
13	45.5	15.5	14.5	15.5
11	44.5	14.5	14	16
9	46.5	14	15.5	17
13	47	16	15	16
12	41.5	14	12.5	15

Case 6

BDI-II Score	Total Alliance Score	Goal Subscale Score	Task Subscale Score	Bond Subscale score
30	48.5	16	16.5	16
21	47	16	16	15
7	45	15	14.5	15.5
19	52	17	17	18
9	46.5	15	16	15.5
6	49	15.5	15	18.5
6	47.5	16.5	15	16
6	53	17.5	18	17.5
16	42	14	12.5	15.5
6	47.5	17	16	14.5

Case 7

BDI-II Score	Total Alliance Score	Goal Subscale Score	Task Subscale Score	Bond Subscale score
36	40.5	10.5	14.5	15.5
33	48.5	16	16	16.5
29	45	11.5	15	18.5
40	52	18	15.5	18.5
32	47	14.5	16	16.5
29	49.5	17	16	16.5
33	45.5	15	14.5	16
38	46.5	13.5	16	17
24	47	14	17.5	15.5
23	39.5	12.5	13.5	13.5

Case 8

BDI-II Score	Total Alliance Score	Goal Subscale Score	Task Subscale Score	Bond Subscale score
52	41	11.5	11	18.5
48	46.5	15.5	15	16
54	35.5	7.5	12	16
53	37.5	11.5	12	14
54	42	13.5	13.5	15
48	46	16	14	16
50	43	10	15.5	17.5
48	45	14.5	15	15.5
46	43.5	15	14	14.5
53	33.5	11	10	12.5

Case 9

BDI-II Score	Total Alliance Score	Goal Subscale Score	Task Subscale Score	Bond Subscale score
28	45.5	14.5	15.5	15.5
30	44.5	16	13.5	15
19	47.5	14.5	16	17
22	44	14.5	14.5	15
20	44.5	14	15	15.5
16	44	14.5	15.5	14
17	40.5	13	14.5	13
16	40.5	11.5	16	13
16	43.5	15	15	13.5
7	36	11	13	12

Case 10

BDI-II Score	Total Alliance Score	Goal Subscale Score	Task Subscale Score	Bond Subscale score
22	38	12.5	13.5	12
19	43	13.5	16	13.5
21	36.5	10	14	12.5
20	38	12.5	12.5	13
16	45	16	15	14
14	35.5	12	13	10.5