The Therapeutic Alliance in Cognitive Therapy for Depression in Combination with Antidepressant Medication: Relations to Subsequent Symptom Change and Treatment Retention

THESIS

Presented in Partial Fulfillment of the Requirements for the Degree Master of Arts in the Graduate School of The Ohio State University

By

Elizabeth T. Ryan, M.A.

*****

The Ohio State University

2010

Master’s Committee:

Dr. Daniel R. Strunk, Advisor

Dr. Jennifer S. Cheavens

Dr. Steven J. Beck
ABSTRACT

While Cognitive Therapy (CT) has been established as an efficacious treatment for depression, less is known about the exact mechanisms by which CT achieves its effect. The therapeutic alliance is one of the most widely studied process variables across treatments and results suggest that there is a small, but consistent positive association between alliance and outcome. The results, however, are less clear when the alliance is studied within the context of CT specifically. Methodological differences may partly explain these discrepant findings, highlighting the need to address the most appropriate methodological design when studying the alliance-outcome relationship. Specifically, including only subsequent change in the analyses ensures that the alliance is not confounded with prior symptom improvement. In the current study, alliance was rated by independent observers during the first three sessions of CT within a sample (N = 176) of patients with depression undergoing a course of CT in combination with anti-depressant medication. The overall alliance was predictive (at the level of a non-significant trend) of subsequent early symptom improvement, with the subscale reflecting agreement on treatment tasks emerging as a significant predictor, as well as significantly predicting the rate of improvement over a 16 week period of treatment. The alliance was not found to be a predictor of drop-out rate across the course of treatment. Overall, results suggest while the alliance may be a significant determinant of symptom change, its effects are small.
Dedicated…

…to my parents, for their unwavering support throughout the years

…to my husband, for taking a chance
ACKNOWLEDGMENTS

I wish to thank my advisor, Daniel Strunk, and my research partner, Andrew Cooper, for their continued support and without whom this project would never have come to fruition. I also wish to thank my committee members, Jennifer Cheavens and Steven Beck, for offering their time and thoughtful recommendations, as well as the many undergraduate research assistants in the Depression Research Laboratory who contributed to data collection.
VITA


2003 ............................................................... B.A., Psychology (magna cum laude), The
College of William and Mary.

2004 – 2007 ................................................ Research Assistant, the Behavioral and
Psychopharmacology Research Laboratory,
McLean Hospital, Belmont, MA.

2006 ............................................................... M.A., Psychology, Boston University.

2007 – 2008 ................................................ University Fellow, The Ohio State University.

2008 – present ............................................. Graduate Teaching Assistant, Department of
Psychology, The Ohio State University.

PUBLICATIONS

1. Lindsey, K.P., Lukas, S.E , MacLean, R.R., Ryan, E.T., Reed, K.R., & Frederick, B.
deB. (2009). Design and validation of an improved nonferrous smoking device for self-
administration of smoked drugs with concurrent fMRI neuroimaging. Clinical
Electroencephalography and Neuroscience, 40, 21-30.

2. Elman, I., Chi, W.H., Gurvits, T.V., Ryan, E.T., Lasko, N.B., Lukas, S.E. & Pitman,

reported speech. Memory, 14, 395-409.

plasma, serum, and whole blood ethanol concentrations: Impact of storage conditions and
collection tubes. Journal of Analytical Toxicology, 32, 505-510.


**FIELDS OF STUDY**

Major Field: Psychology
**TABLE OF CONTENTS**

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abstract</td>
<td>ii</td>
</tr>
<tr>
<td>Dedication</td>
<td>iii</td>
</tr>
<tr>
<td>Acknowledgements</td>
<td>iv</td>
</tr>
<tr>
<td>Vita</td>
<td>v</td>
</tr>
<tr>
<td>List of Tables</td>
<td>x</td>
</tr>
<tr>
<td>List of Figures</td>
<td>xi</td>
</tr>
</tbody>
</table>

Chapters

1. Introduction.................................................................................................................... 1
   Alliance as Predictor of Outcome ................................................................. 1
   Methodological Challenges to Address ............................................................. 3
   Timing of Measurement ....................................................................................... 4
   Temporal Precedence ......................................................................................... 5
   Alliance and Drop-out ....................................................................................... 13
   Cognitive Therapy and Pharmacotherapy for Depression .................................. 14
   Problems with Definition and Measurement of Alliance .................................... 16
   Choosing an Alliance Scale ............................................................................... 16
   Method of Data Collection ............................................................................... 20
Hypothesis 1: Analysis of Session-to-Session Symptom Change .................33
Hypothesis 2: Analysis of Drop-out ..................................................................34
Hypothesis 3: Analysis of Longer-term Symptom Change .........................35
Secondary Analyses .............................................................................................36

4. Discussion ........................................................................................................37
   Early Symptom Change .........................................................................................38
   Treatment Retention..............................................................................................41
   Long-term Symptom Change ...............................................................................41
   Limitations ........................................................................................................42
   Conclusions ..........................................................................................................46

List of References ..................................................................................................47
Footnotes ................................................................................................................53
Appendix A: Tables ...............................................................................................54
Appendix B: Figures ...............................................................................................62
## LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Studies of the Alliance-Outcome Association with Temporal Considerations</td>
<td>55</td>
</tr>
<tr>
<td>2</td>
<td>Average BDI-II Scores Across Sessions</td>
<td>56</td>
</tr>
<tr>
<td>3</td>
<td>Correlations between the Mean Alliance for Individual Patients</td>
<td>57</td>
</tr>
<tr>
<td>4</td>
<td>Association between Alliance and Improvement in the Following Session</td>
<td>58</td>
</tr>
<tr>
<td>5</td>
<td>Average Alliance as Predictor of Drop-out</td>
<td>59</td>
</tr>
<tr>
<td>6</td>
<td>Alliance as Predictor of HRSD Slope Following Session 3 Through 16</td>
<td>60</td>
</tr>
<tr>
<td>7</td>
<td>Association between Symptoms and Alliance in the Following Session</td>
<td>61</td>
</tr>
</tbody>
</table>
LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Change Scores in BDI-II over Early Between-Session Intervals</td>
<td>63</td>
</tr>
<tr>
<td>2</td>
<td>Length of Time in Treatment</td>
<td>64</td>
</tr>
</tbody>
</table>
CHAPTER 1

INTRODUCTION

Cognitive Therapy (CT) has repeatedly been shown to be an efficacious treatment for Major Depressive Disorder (MDD; for a review, see Strunk & DeRubeis, 2001). Recent studies have established that CT is at least as efficacious as pharmacotherapy over the course of one year (DeRubeis, Gelfand, Tang, & Simons, 1999; DeRubeis et al., 2005) and provides more enduring effects than discontinued medication (Hollon et al., 2005). However, the exact mechanisms by which CT achieves its effect have not yet been clearly identified. The therapeutic alliance is arguably the most frequently studied psychotherapy process variable in the literature – between the years 1977 and 2000, over 2,000 publications on this topic have been located (Horvath & Bedi, 2002) – but the support for its specific role in treatment is varied. In this paper, I attempt to extend and improve upon the previous findings in the alliance-outcome literature as outlined below.

The alliance as a predictor of outcome

The alliance has repeatedly been found to be correlated with outcome across treatments; in two meta-analyses conducted on the alliance-outcome association (Horvath & Symonds, 1991; Martin, Garske, & Davis, 2000), the average effect sizes were $r = 0.26$ and $r = 0.22$, respectively, which is a reliable, though small association. The majority of studies to date have been conducted with treatments based on humanistic and psychodynamic perspectives (Constantino et al., 2002), but there has been a surge of
studies in the past two decades attempting to study the alliance in the context of more structured therapeutic approaches. The findings from these studies have largely been similarly positive. Castonguay, Goldfried, Wiser, Raue, and Hayes (1996), for example, looked at a sample 64 patients with MDD who underwent either CT, CT plus anti-depressant medication (CT+ADM), or ADM plus clinical management (ADM+CM) and found that the alliance was significantly related to improvements in depressive symptoms and overall global functioning in the conditions which included CT. Similarly, Hardy Cahill, Shapiro, Barkham, Rees, and Macaskill (2001) found a positive association between alliance and outcome for 24 MDD patients in CT. These findings show support for the notion that alliance can predict positive changes in a structured treatment such as CT.

Results have not always been supportive of the alliance-outcome association in CT, however. Krupnick and colleagues (1996) found that patient contributions to the alliance (and not therapist contributions) were a significant predictor of change in depressive symptoms with a large sample of 225 MDD patients. However, there was also a significant interaction across treatments that signified early alliance within CT specifically was less strongly predictive of outcome (as measured by the BDI), as compared to the interpersonal therapy (IPT), ADM, and Placebo (PLA) groups. Raue, Goldfried and Barkham (1997) also compared CT and IPT in a sample of 57 people diagnosed with MDD and found that while the alliance scores were higher for patients in the CT group, there were no treatment interactions, and that sessions rated as “high impact” and “smooth” were characterized with higher alliance scores regardless of
whether the client was in CT or IPT. These findings suggest that perhaps the alliance is not especially important in determining a positive outcome within CT specifically.

The discrepancies across the research of the alliance-outcome association within structured treatments such as CT raises the need to address some of the methodological problems that potentially impacted the results; until certain issues are addressed, it will be impossible to determine the role that the alliance plays in the process of change within CT, or any other form of psychotherapy.

Methodological challenges to address

In non-experimental studies examining the relationships between the alliance and outcome, there are three circumstances which might account for an association between the alliance and outcome: (1) alliance causes outcome; (2) changes in outcome cause alliance; or (3) a third variable accounts for the association of alliance and outcome.

Unfortunately, it is impossible to account for and rule out all potential confounds to the alliance-outcome relationship when studying it in the context of a non-experimentally controlled trial where the alliance is merely measured, not manipulated. Steps can be taken, however, to rule out reverse causality, which in this case refers to changes in outcomes causing changes in alliance. First, the timing of the measurement can be controlled in an attempt to reduce the chance of a resulting in an erroneous relationship. Second, the analyses need to be conducted carefully so that only subsequent symptom change is being predicted, ensuring that any prior change does not contribute to (and perhaps artificially inflate) the alliance-outcome association.
Timing of the measurement

One issue that needs to be taken into consideration when investigating the direction of the alliance-outcome relationship is the timing of the alliance measure. In order to capture the alliance as a predictor of the most therapeutic change possible, it is important to focus on measuring the construct early in treatment, when a substantial amount of symptom improvement occurs (Ilardi & Craighead, 1994). Few researchers have recognized the importance of the chronology of the alliance measure and, as a result, do not specify why they have chosen to run measures at certain times during treatment – some proclaim that it was random selection (e.g., Castonguay et al., 1996) while others leave the question unanswered (e.g., Hardy et al., 2001). This creates a problem of variability across studies, making it difficult to equate or compare results, particularly if the alliance is measured subsequent to extensive symptom change. Further, some studies (e.g., Kivlighan & Shaunnessy, 1995) have observed small and nonsignificant correlations between therapist and patient alliance ratings early in treatment, but in later sessions correlations became moderate to large, and were only then found to be related to outcome. One explanation offered for this phenomenon is that the patient and therapist grew to share a common perception of the working alliance (Kivlighan & Shaunnessy, 1995). Another explanation for this finding, however, is not that the therapist and client gained a mutual understanding of the relationship but that the relationship improved as the client’s symptoms improved (DeRubeis & Feeley, 1990).

Research has supported this concept of reverse causality, the notion that clinical improvement causes a positive alliance. For example, DeRubeis and Feeley (1990) found that alliance later in treatment was predicted by prior symptom reduction. In other words,
as a patient begins to feel better, it is possible that it becomes easier to bond with the therapist, to feel more agreeable, and to have the energy and motivation to follow through on self-improving tasks, such as homework assignments. Thus there is evidence that prior improvement of symptoms should be taken into account if the causal direction of the relationship between the alliance and outcome is to be determined. Therefore, when designing an alliance-outcome study, it is prudent to use methods that do not allow for reverse causality as an explanation for the effects obtained.

**Taking temporal precedence into account**

It is important to carefully design methodology to most appropriately address the specific research questions, as variability in methods appears to lead to discrepant results, as can be seen in the case of Feeley, DeRubeis and Gelfand (1999) and Castonguay and colleagues (1996). The findings from both of these papers were obtained from the same patient sample, but reached different conclusions – Castonguay and colleagues found a positive alliance-outcome association, whereas Feeley and colleagues failed to find a significant relationship. This is likely attributable to differences in the measurement of the alliance and symptom change. First, alliance measurements were obtained at different time points – Feeley and colleagues measured alliance during the second therapy session from each of the patients to then predict outcome, whereas Castonguay and colleagues measured the alliance at either session 4, 5, 6 or 7 to predict outcome. The second difference was how outcome was measured. Feeley and colleagues examined change in BDI scores from the session subsequent to the alliance measure through to the end of acute treatment. In contrast, Castonguay and colleagues examined change in symptoms through the end of acute treatment but, while they did control for initial symptom
severity, they did not limit their analyses to changes that occurred only subsequent to the measurement of the alliance. It is possible that these methodological differences could partly account for the discrepant conclusions – Castonguay and colleagues concluded that there was a significant positive association between the alliance and the outcome measure, whereas Feeley and colleagues found not only that the early alliance was not associated with outcome, but that the later alliance scores were actually predicted by early symptom change. Evidence suggests that the alliance-outcome correlation predicted by Castonguay and colleagues may have erroneously been attributed to the strength of the alliance, but instead it was merely a reflection of differences in symptoms levels, and potentially previous therapeutic gains, that were not accounted for. This implies that researchers should attend to the temporal precedence of the alliance measure and control for current symptom levels to ward off differences in clinical improvement when measuring the effect of process variables on outcome.

A number of studies have addressed the need for temporal consideration when measuring the effects of alliance on outcome, but results have been mixed (see Table 1 for a summary). Several studies failed to find a significant relationship between alliance and subsequent symptom improvement. DeRubeis and Feeley (1990) looked at a sample of 25 patients who received 12 weeks of CT for a primary diagnosis of MDD, measuring alliance at session 2 and a session each within each the last three quadrants of treatment. They did not find a significant association between early alliance and outcome when measuring subsequent symptom change ($r = .10$). As mentioned previously, they also found that in the last half of treatment, the alliance was greater for clients that had experienced greater symptom improvement ($r = .48$), supporting the notion of reverse
causality of the alliance-outcome association (i.e., that it is clinical improvements that are predicting a good alliance). Similarly, Feeley and colleagues (1999) examined a group of 25 patients diagnosed with MDD who underwent 12 weeks of either CT as a monotherapy (n = 12) or CT in combination with pharmacotherapy (n = 13). Alliance was measured by independent observers at session 2 and a session each within the last two quadrants of the 12 week treatment. They also failed to find a relationship between early alliance and subsequent symptom (although there was a non-significant trend and an effect size of $r = .27$).

While the sample sizes in these two studies were small, partial support for their conclusions was found in a larger sample by Barber and colleagues (1999), who studied a group of 252 patients seeking treatment for cocaine dependence, a potentially very different sample than patients with MDD. Patients were randomized to either supportive expressive therapy (SE), CT, brief psychodynamic therapy (B-PT), or drug counseling based on the 12-step model, and the alliance was measured at sessions 2 and 5. The alliance was not predictive of changes in their primary measures of interest (addiction severity and global severity), but, at first glance, it appeared that the alliance at both sessions 2 and 5 did predict changes in BDI scores at 1-month and 6-month assessments. In fact, the alliance at session 2 also predicted treatment retention, particularly for older patients. Barber and colleagues also found that if the alliance was rated high early in treatment, patients who stayed in the study reported a better outcome and if the early alliance was rated low, patients who stayed in the study reported worse outcomes. This suggests that if a proper alliance is not formed at the beginning of treatment, then it could be detrimental to the outcome. These results appear to be promising, but one must
interpret them with caution, as when the authors controlled for prior symptom change, they did not find a significant relationship between alliance and symptom improvement (reduction in drug use and global severity, $r = .01$). BDI, however, was not included in this last step of analyses, so controlling for previous improvement in depressive symptoms may or may not have had had an impact on their positive findings.

Another recent study failed to find a significant alliance-outcome association. Strunk, Brotman and DeRubeis (2009) examined the first four sessions of CT for 60 adult patients with MDD. In this regard, the alliance was captured beyond the typical “snapshot” assessment, which is taken at one time point (or at various points with substantial amounts of time in between measures), increasing the precision of the measurement. They applied a statistical model to capture session-by-session change after controlling for current symptom severity during early treatment, when symptom improvement occurs most rapidly. The alliance was not found to significantly predict short-term improvements, nor did it predict outcome at the end of treatment. The effect size ($r = .15$) for the session-by-session analyses, however, was comparable to the studies that did find a statistically significant relationship.

Despite these mixed findings, several studies have found significant alliance-outcome relationships when subsequent symptom change was measured. Barber and colleagues (2000) completed a study with 86 patients who underwent 16-52 sessions of SE for diagnoses of either Generalized Anxiety Disorder (GAD; $n = 44$), MDD ($n = 11$), Avoidant Personality Disorder ($n = 19$), or Obsessive-Compulsive Personality Disorder ($n = 14$). They measured alliance at session 2, 5, and 10 and, after controlling for intake BDI, they found that the alliance at all three sessions predicted subsequent change at both
the 4-month assessment point and at termination. Similar to DeRubeis and Feeley (1990),
they also found that alliance measured later in treatment was associated with prior
changes in depressive symptoms. These findings suggest that the alliance does predict
changes in outcome in SE for a variety of disorders, but that improvement in outcome
also predicts a stronger alliance later in treatment.

In a recent study, Zuroff and Blatt (2006) analyzed data from the National
Institute of Mental Health Treatment for Depression Collaborative Research Project
(TDCRP), which was comprised of 239 patients randomly assigned to 16 weeks of either
interpersonal therapy (IPT), CT, ADM plus clinical management (ADM+CM), or
placebo plus clinical management (PLA-CM) for treatment of MDD. Of these 239,
Zuroff and Blatt included in their analyses the 191 patients who had complete data for the
alliance and early ratings of maladjustment, as well as their other variable of interest, the
patients’ perceptions of the quality of the therapeutic relationship, which taps into
whether the client feels that the therapist is providing enough empathy and whether the
client is being open and honest with the therapist. They found that a perceived positive
therapeutic relationship early in treatment predicted not only the slope of subsequent
symptom improvement (controlling for prior symptom change), but also predicted better
adjustment at an 18-month follow-up assessment ($r = .15$). There was no significant
interaction with treatment, suggesting that fostering a strong, positive early alliance is
important not just within specific technical psychosocial treatments, but apparently has
positive implications for people seeking clinical management of pharmacotherapy
(including placebo treatment), which in this case was prohibited from utilizing any
specific psychotherapeutic techniques.
The largest study to date exploring the alliance-outcome relationship also found positive results. Klein and colleagues (2003) completed a study with 367 patients seeking treatment for MDD who were randomly assigned to either Cognitive-Behavioral Analysis System of Psychotherapy (CBASP; McCullough, 2000) as a monotherapy (n = 169) or CBASP in combination with ADM (CBASP+ADM, n = 198). Alliance was measured by patient reports of the alliance for sessions 2, 6 and 12; after controlling for baseline severity, the alliance measured at session 2 was found to predict subsequent changes from weeks 3 through 12 ($r = .14$). These results held true even when controlling for other potential third variables, such as gender, chronicity, comorbid anxiety, substance use, personality disorders, level of social functioning and a history of childhood abuse. As with Zuroff and Blatt (2006), treatment condition did not moderate this effect, although the combination treatment group did report higher ratings of early alliance. This study serves as the strongest evidence for an alliance-outcome association given its statistical power, the number of controlled elements, and its consideration of the crucial temporal precedent.

Taken together, many questions still remain as to the precise role that the alliance plays in treatment with regards to predicting levels of improvement. The negative findings from the studies that had temporal consideration (DeRubeis & Feeley, 1990; Barber et al., 1999; Feeley et al., 1999, Strunk et al., 2009) vary greatly in meaningful ways. For one, the populations differ – DeRubeis and Feeley, Feeley and colleagues, and Strunk and colleagues all studied patients with MDD, while Barber and colleagues concentrated on people with cocaine dependence, therefore they predominantly measured outcome as a change in drug use as opposed to a reduction in depressive symptoms.
Despite the emphasis on a depressed population within the negative findings, however, there are methodological differences that make drawing substantial conclusions difficult. First, the treatment conditions are not the same – all three MDD studies utilize CT in some capacity, but approximately half of the patients in the study by Feeley and colleagues were also receiving concurrent pharmacotherapy and these conditions were not separated in the analyses. Second, the sample sizes are quite small in two of the three studies, particularly if one tries to look specifically at MDD patients receiving CT – both DeRubeis and Feeley, and Feeley and colleagues have a sample size of 25 patients, and that total includes both treatment conditions in the latter study; only Strunk and colleagues had a sample of reasonable size (n = 60) to have enough power to establish confidence in an accurate effect. Methodologically, the study by Strunk and colleagues also differs from the other studies, as it utilizes session-by-session ratings of the alliance with the intention of increasing the precision of catching rapid symptom improvement early in treatment, as opposed to utilizing endpoint data.

The studies with temporal considerations that did find a positive relationship between the alliance and outcome (Barber et al., 2000; Klein et al., 2003; Zuroff & Blatt, 2006), however, are also inconsistent. First, all three studies do include an MDD population within their sample, but this group is a minority in the overall sample collected by Barber and colleagues (MDD, n = 11; GAD, n = 44; Avoidant Personality Disorder, n = 19; Obsessive-Compulsive Personality Disorder, n = 14). In addition, CT for MDD is only represented in part by Zuroff and Blatt (one of four conditions, including IPT, ADM+CM, and PLA-CM). The treatment used in the Klein and colleagues paper, CBASP, is based in part on cognitive and behavioral principles, but
strays from the established manualized treatment and integrates more of a focus on interpersonal issues. This makes it difficult to generalize the findings to CT, especially when considering an interpersonal measure such as the alliance. Second, the sample size is no longer a problem upon first glance, as they range from 86 (Barber et al., 2000) to 161 (Zuroff & Blatt, 2006) to 367 (Klein et al, 2003). But, again, analyses include heterogeneous samples and reduces the sample size and predictive power within each treatment condition of interest, in the case of Zuroff and Blatt (161 patients within four treatment conditions), and within the population of interest, particularly in the case of Barber and colleagues (86 patients divided into four diagnostic categories, primarily GAD). Third, the “alliance” measured in the study conducted by Zuroff and Blatt included only the patient contributions to the therapeutic relationship and eliminated the therapist contributions altogether. This decision was predicated on the findings in the study by Krupnick and colleagues (1996), which suggested that patient contributions of alliance were more predictive, as they failed to find therapist contributions to alliance to be a significant predictor outcome. However, by limiting the alliance measure to only the patient perception of and input on the relationship, they ignore an instrumental component in developing the working relationship (i.e., therapist contribution) and make it difficult to compare the resulting findings with other studies of the alliance. Thus, the conceptual and methodological issues complicate the interpretations of both the positive and negative findings with the alliance-outcome association in the literature.

Alliance and Drop-out

There is also interest in the alliance as a predictor not only of clinical outcome, but as a predictor of retention in treatment. One of the major challenges across treatments
for mental illness is the high level of patients dropping out of therapy, with approximately 35–47% of patients ending treatment prematurely (Wierzbicki & Pekarik, 1993) and greater than 40% of patients dropping out of treatment with ADM alone within three months (adequate course of treatment defined as six months; Venturini, Sung, Nichol & Sellner, 1999). Despite the prevalence of premature termination, there are relatively few studies that focus on predictors of treatment drop-out, and even fewer investigating the role of process variables such as the therapeutic alliance on patient retention.

Most of the studies that have tackled this issue yielded findings that suggest that the alliance is an important predictor of drop-out rates. Two separate studies completed by Samstag and colleagues (Samstag, Batchelder, Muran, Safran, & Winston, 1998; Samstag, Muran, Wachtel, Slade, Safran, & Winston, 2008), for example, investigated samples of patients who were primarily diagnosed with MDD and underwent one of five time-limited, manualized treatments (two brief psychodynamic, CBT, SE, and relational). The investigators broke patients down into one of three categories: drop-outs, completers with poor outcome and completers with good outcomes. The results from both of the studies concluded that patients who dropped out had poorer alliances with their therapists than either of the groups that completed the full course of treatment. Unfortunately the sizes of the groups in each treatment type were small, so the analyses were collapsed across treatments and there was no additional information reported on the role of alliance specifically in the CT condition.

Arnow and colleagues (2007), who utilized the same dataset as Klein and colleagues (2003) mentioned above, found that chronically depressed patients in either
the cognitive-based psychotherapy CBASP or a combination of CBASP and ADM (CBASP+ADM) who terminated prematurely reported significantly lower alliances early in treatment ($F = 6.9, p = .009$). They also found that, even though drop-out rates were equivalent across treatment conditions, patients who dropped out of the CBASP+ADM condition remained in treatment significantly longer than patients in either of the other two treatment groups (CBASP+ADM mean = 40 days vs. CBASP mean = 28 days, ADM mean = 27 days). Those patients who did drop-out of the CBASP+ADM condition were also less likely to attribute medication side effects as a reason for ending treatment early, as opposed to those patients who dropped-out of the ADM alone condition. These findings are particularly noteworthy as they are directly applicable to the current research design, which focuses on a cognitively-based psychotherapy in combination with ADM.

*CT and pharmacotherapy for MDD*

While all of the participants in the current study received CT for MDD, they also were all treated with concurrent pharmacotherapy. The combination of CT with antidepressant medication has repeatedly been shown to be as efficacious as either treatment provided as a monotherapy and, in multiple cases, the CT+ADM conditions have numerically outperformed the monotherapies, although often the effects are small and do not achieve statistical significance (Hollon, DeRubeis, Evans, Wiemer, Garvey, et al., 1992; Friedman, Detweiler-Bedell, Leventhal, Horne, Keitner, & Miller, 2004). This combination of treatment is especially important to study because, according to two national surveys, 45% of people who seek treatment for MDD as outpatients receive both psychotherapy and pharmacotherapy (Olfson et al., 2002). Therefore there are crucial real-world implications regarding when, if, and how combination treatment achieves its
efficacy and unfortunately there is an impoverished literature base, particularly within the realm of process research in this context (Otto et al., 2005).

Several studies have found a significant advantage from the addition of ADM to a course of cognitively oriented psychotherapy, in both inpatient (Bowers, 1990) and outpatient samples (Blackburn, Bishop, Glen, Whalley & Christie, 1981). Specifically, Keller and colleagues (2000) analyzed data from the same sample of patients that Klein and colleagues (2003) used above and found that patients who underwent a combination of CBASP+ADM had a significantly higher response rate (73%) than either treatment on its own (48%). The combination condition also sustained remission criteria longer than either the CBASP or ADM conditions administered on their own (48% vs. 33% and 29%, respectively).

Perhaps most importantly, combination treatments have been found to better prevent relapse over time compared to ADM alone. When the treatments were provided in a complementary fashion during acute treatment, researchers have found that people in the combination treatment had significantly lower rates of relapse than ADM alone (28% vs. 66%, respectively; Simons et al, 1986) at both one-year follow ups and at two-year follow-ups, where participants in either the CT or combination conditions had lower relapse than those in short-term and long-term ADM (20% and 15% vs. 32% and 50%, respectively; Evans et al., 1992). When CT was provided after a successful course of treatment with ADM, patients relapsed significantly less at four years than those who were provided with continued clinical management (CM; 35% vs. 70%) and, despite the lack of difference in relapse rate at six years after treatment, people who underwent CT
were shown to have fewer new episodes of depression than those in CM (Fava et al., 1998).

While this evidence is promising, more research needs to be conducted in the context of combination treatment on the specific mechanisms of change through which CT achieves its efficacy and how the most efficacious and cost-effective forms of treatment can continue to fostered. The incorporation of CT into a treatment plan for MDD may be a favorable choice because it is generally more tolerable (i.e., no aversive side effect profiles) and more cost-efficient because of its short-term nature and ability to prevent relapse, therefore it appears to have overall better long-term outcomes (Otto et al., 2005).

**Problems with the definition and measurement of the alliance**

*Choosing an alliance scale*

Another nontrivial issue that needs to be dealt with in research with alliance as a process measure is the discrepancies between defining and measuring the construct. The term “alliance” is used to describe a number of related constructs, but at this point there is not a universally accepted definition (Horvath & Bedi, 2002). The strength of the relationship between therapist and patient has long been considered an important component of psychotherapy, and arguably the most influential conception of the therapeutic relationship was introduced by Rogers’s (1957) and client-centered therapy. According to this approach, the therapeutic relationship is comprised of accurate empathy, genuineness, and unconditional positive regard (Rogers, 1957). According to such humanistic and psychodynamic orientations, the alliance plays a critical role in establishing change in treatment by providing patients with a stable relationship and
allowing the patient to incorporate the nonjudgmental approach of the therapist into their own self-image (Meissner, 2007). Cognitive and behavioral therapies, on the other hand, have often been criticized for placing less emphasis on the alliance in lieu of a focus on building skills and problem solving in order to keep the patient in the here-and-now.

The “working alliance” is a model of the therapeutic relationship that breaks the construct down into three components: (1) an agreement between client and therapist about the goals of therapy, (2) an agreement on the steps that need to be taken to meet these goals, and (3) the personal bond that develops between the client and therapist (Bordin, 1979). Conceptually, this definition of alliance seems more applicable to CT, which emphasizes an active and collaborative partnership, as opposed to the strong focus on the affective quality of the relationship that is seen in many psychodynamic and humanistic treatment orientations. Therefore, it appears logical that when investigating the alliance in CT, one would choose a measure that captures all of the intended domains of the alliance defined in a theoretically appropriate fashion.

In the two recent meta-analyses of the alliance-outcome relationship (Horvath & Symonds, 1991; Martin et al., 2000), almost two dozen different alliance measures were used to collect data, including several ambiguous categories the meta-analysts called “other,” “non-standardized scale,” and “various inpatient therapeutic alliance scales.” Both studies, however, agreed upon the four most frequently used scales: the Penn Helping Alliance Questionnaire (HAq; Luborsky et al., 1985); the Vanderbilt Therapeutic Alliance Scale (VTAS, Suh, O’Malley & Strupp, 1986); the California Therapeutic Alliance Rating Scale (CALPAS; Marmar, Weiss & Gaston, 1989); and the Working Alliance Inventory (WAI; Horvath & Greenberg, 1989). Each of these four measures
operationalize the alliance a little differently and emphasize different aspects of the relationship: HAQ focuses on the balance between the compassionate and patient collaboration features of alliance (Luborsky et al., 1983), the VTAS considers both the positive and negative characteristics of the client and therapist (Suh et al., 1986), the CALPAS highlights not only the cooperation and goal setting, but also the positive and negative contributions of both the therapist and the patient (Marmar et al., 1989), and the WAI emphasizes agreement and collaboration on specific therapeutic tasks and goals (Horvath & Greenberg, 1989).

Of these four scales, the WAI was used the most frequently in the literature. As the frequency of use speaks to the popularity of the scale, then the WAI comes across as a well-used, and therefore widely comparable, measure of the alliance. While many of the earlier alliance scales (including the HAq and the VTAS) were created within a general psychodynamic framework of the alliance (i.e., focuses on the affective quality of the relationship), the WAI is a scale that was designed as a transtheoretical measure of alliance with the intention of establishing an explicit connection between a measurable construct of alliance and a wide-ranging theory of therapeutic change. Specifically, the WAI aims to encapsulate Bordin’s (1979) conception of the alliance as three components: the therapeutic Bond, measuring mutual liking and attachment; the collaborative setting of Goals addresses the extent to which the goals of therapy are important, mutually defined and achievable; and the agreement on Tasks needed to undertake in order to improve the client’s current situation and accomplish the therapy goals.

The WAI is not the only alliance scale created to span theoretical orientations, however. The CALPAS was also designed to be a broad-spectrum measure of alliance,
and it is made up of four expansive subscales: Therapist Understanding and Involvement, the therapist contribution component; Patient Commitment, referred to as the working alliance; Working Strategy Consensus, the patient-therapist agreement on treatment goals; and Patient Working Capacity, the element defined as the therapeutic alliance. In theory, the CALPAS appears to be an appropriate scale to use in CT research, as it captures the active working relationship, the focus on goals, and the importance of an affective bond. However, in order to choose an alliance measure that will be used in a study designed with temporal considerations, a key concern is that the scale does not attempt to capture symptom change, as this would confound the necessary control for prior change. Unfortunately, the Patient Commitment subscale of the CALPAS is contaminated with questions aiming to assess the patient’s satisfaction with treatment and the outcome thus far. Therefore, the CALPAS does not appear to be the most appropriate measure of the alliance in studies with temporal precedence.

In addition to being a popular measure and conceptually relevant to the context of CT, the WAI also achieved the best inter-rater reliability of all the alliance measures (Martin et al., 2000). Despite the differences among these measures, however, several studies have looked at the strength of the correlations between the total scores on the four measures (e.g., Fenton et al., 2001, Hatcher & Barends, 1996, Tichenor & Hill, 1989), and found medium to high intercorrelations between measures. This suggests that while the measures may capture slightly different aspects of the alliance, connections can still be drawn regarding the “alliance” across studies with varying use of these four well-used measured.
**Method of data collection**

Beyond the variations in the creation of explicit definitions and measurements of the alliance, however, there are also problems with the disparity of data collection. Three out of the four most used measures can be utilized by patients, therapists or independent observers; this flexibility and depth of assessment may seem to be a benefit, but results have suggested that the measurement of the alliance differs depending on the rater’s point of view. Some studies (e.g., Barber et al., 1999) have utilized more than one measure of alliance from both the patient and therapist perspective and found that the scales not only measured different aspects of the alliance, but that the correlations were often low between patient and therapist ratings. Patient ratings of alliance have also been found to be almost as predictive of outcome as those of independent observers, and both are more predictive than therapist ratings (Horvath & Symonds, 1991), although the meta-analyses have not limited the studies included to those with temporal precedence of the predictor to outcome. This discrepancy, particularly between therapist and patient ratings, appears to be greatest when rated in early sessions (Horvath & Bedi, 2002). Horvath and Bedi have suggested that it is likely that the conception of the alliance is especially dissimilar early in treatment because the therapist and patient are focused on different aspects of therapy; the client is often looking for someone to whom they can feel safe with and connected to, as they are about to embark on a very personal and difficult experience, whereas the therapist is likely evaluating the alliance from the perspective of how therapeutic response should appear within their own particular theoretical orientation (Horvath & Bedi, 2002).
While asking patients to complete alliance measures may slightly increase their burden of providing repeated self-reports, practically it makes sense to utilize patient and therapist ratings because it eliminates the profusion of time that is demanded in order to first train independent observers and then to code data from audio/video recordings of the therapy sessions. Unfortunately, relying on self-report measures bears a number of problems, including, but not limited to this issue of differing conceptions of the alliance (Elvins & Green, 08). When therapists are rating the alliance, they are likely influenced by the expectations of therapeutic success and the definition of “alliance” within their adopted theoretical frameworks, which, as previously discussed, is rout with problems because of the lack of a universal definition of the construct. This idea is supported by the finding that there is much more variation found within therapist-rated scores that those rated by the patient (Kivlighan & Shaughnessy, 1995).

The problems with relying on the patient perspective, however, are similarly complicated. Certain subjective elements of the patients’ experience may influence their alliance ratings, including insight into their own involvement in treatment, personal levels of comfort, feelings of progress that they are or are not making in treatment (Eugster & Wampold, 1996), or prior expectations (Constantino et al., 2005). The perspectives of the therapists and patients, therefore, appear to be inherently different and perhaps irreconcilable without proper training and provision of explicit definitions for the properties of the alliance that are intended to be measured. Utilizing independent observers may miss some of the subjective components of the therapeutic relationship, such as attitude and internal motivation, but it allows for careful training to capture the
same operationalization of alliance and avoids the multiple confounds found with therapist and patient ratings.

**The current study**

The current study aims to address some important gaps in the alliance-outcome literature by addressing several important common limitations of previous research.

*Ruling out Reverse Causality*

Arguably the most important aspect of the current study design is addressing the issue of temporal precedence. Analyses will explicitly focus on examining only subsequent symptom change to avoid the possibility of reverse causation. In other words, by controlling for current symptom levels in the model, the findings cannot suggest that previous clinical improvement is aiding in the formation of a good alliance. For longer-term analyses, HRSD scores, the primary outcome measure, were compiled after session 3 (i.e., after the last alliance scores were rated) through week 16; this ensures that measurement occurs within an amount of time consistent with other process research and studies of the efficacy of time-limited CT (e.g., Hollon et al., 1992; DeRubeis et al., 2005).

The focus of this study is also on early treatment sessions for several reasons. First, there is evidence that a considerable degree of the gains made in psychotherapy (and particularly in CT) occur in the early part of treatment (Ilardi & Craighead, 1994). As such, it may be beneficial to evaluate process measures prior to (and possibly during) this period of rapid change, rather than in later sessions when change may plateau and there are less variations in symptoms that can be captured. Second, using these sessions, in conjunction with controlling for prior symptom improvement, allows assurance that
any effects observed are unidirectional; that is, it addresses the temporal precedence issue highlighted earlier – later sessions may be contaminated by causal relationships which have played out earlier in treatment. Third, the rating of three sequential early sessions creates a more precise measurement of the construct and opens up the opportunity to capture the impact that the developing alliance may have on early symptom gains. Fourth, treatment length in this study is variable, which may preclude examination of specific later sessions for some patients.

**Definition and measurement of the alliance**

The current study is looking to explore the relationship of the alliance in the context of CT for MDD. Because of the emphasis that CT places on structure and collaboration, a conceptually appropriate definition of alliance is that of Bordin’s (1979) working alliance, which captures the agreement on the overall goals of treatment over time, the understanding of the specific tasks that need to be addressed from session to session in order to achieve those goals, and the quality of the personal bond that forms between the therapist and patient. In order to measure this concept of the alliance, the WAI was chosen, primarily because it is not embedded in one theoretical orientation, it is the most widely used measure, it has the highest inter-rater reliability, and the item content does not refer to symptom improvement (Martin et al., 2000). The observer-rated version of the WAI was also utilized, which has found to be a slightly better predictor of outcome than the patient-rated and significantly better than the therapist-rated in a sample of studies with mixed methodologies (Horvath & Symonds, 1991). By employing independent raters, the biases associated with self-report measures, such as social desirability, can be bypassed and the coders can be thoroughly trained to capture the
specific aspects of the alliance within the provided construct definition. Patient ratings may also reflect their cumulative experience to date, whereas observers are only exposed to a single session per patient because sequential ratings might lead to biases, as raters could perhaps get more information about the degree to which a patient has experienced symptom change.

_Hypotheses_

Based on the recent findings of Stunk and colleagues (2009), the only other study that focused on early improvement in a session-by-session design, it is hypothesized that the alliance will come out as a significant predictor of early symptom change; while Strunk and colleagues did not find a significant relationship between alliance and early session-to-session change, the effect size was similar to significant effects found in larger studies (Klein et al., 2003; Zuroff & Blatt, 2006) and the current study has a larger sample size and greater power to achieve a significant effect. It is also hypothesized that the mean overall early alliance will be significantly related to treatment retention, in that a poorer alliance will predict a higher likelihood to drop out. Finally, it is predicted that the alliance will not be able to predict the slope of symptom change over the course of 16 weeks of treatments.
CHAPTER 2

METHODS

Participants

Participants were 176 of 227 patients (78%) in the CT combined with ADM (CT+ADM) condition of a clinical trial comparing CT+ADM to CT alone. The sample was limited to those participants in the CT+ADM for whom there was at least one usable video recording from one or more of the first three CT sessions available. Of the 176 patients in the current sample, 58% were women, 85% were White (9% were African American and 7% were other ethnicities), and the median annual income level was between 30 and 40 thousand dollars. The average age was 43.8 (SD = 13.0). The average age of onset of MDD was 25.2 years old (SD = 13.4), 49.4% reported experiencing three or more episodes, and 31.2% reported suffering from chronic MDD.

All patients included in the study were adults age 18 or older, with a primary current diagnosis of Major Depressive Disorder (MDD), recurrent or chronic, as measured by the Structured Clinical Interview for DSM-IV (First, Spitzer, Gibbon, & Williams, 2001). Participants were excluded from the study if they met any of the following criteria: a history of bipolar disorder; history of psychosis; current non-psychotic Axis I disorder if it constituted the predominant aspect of the clinical presentation; history of substance dependence in the past six months; antisocial,
borderline, or schizotypal personality disorder; subnormal intellectual potential (IQ ≤ 80); evidence of any medical disorder or condition (including pregnancy or risk of pregnancy) that could cause depression or preclude the use of study treatments; current treatment with catecholaminergic antihypertensive medication (including reserpine, beta-blockers, clonidine, alphamethyldopa, etc.); clear indication of secondary gain (e.g., court ordered treatment or compensation issues); or current suicide risk sufficient to preclude treatment on an outpatient basis.

Study Design

The study from which participants were drawn is a three-site clinical trial aimed at evaluating whether CT+ADM reduces risk of relapse and recurrence as compared to ADM alone. The trial was conducted at the University of Pennsylvania in Philadelphia, PA, Vanderbilt University in Nashville, TN, and Rush-Presbyterian-St. Luke’s Medical Center in Chicago, IL. As all analyses focus on patients’ initial response to treatment and analyses are limited to the first year of treatment, only details relevant to the first phase of the trial are described.

Patients in the trial progressed through three phases: an acute treatment phase, a maintenance phase and a continuation phase. The acute treatment phase, the only phase of interest for this study, lasted up to 52 weeks; in order to achieve remission criteria (and thus move from the acute to the maintenance phase of treatment), it was required that patients did not meet DSM-IV criteria for MDD for at least one month and that their symptom severity was sufficiently low (i.e., a score on the Hamilton Rating Scale for Depression (HRSD; Hamilton, 1906) of 8 or less) for a period of one month.
Therapists

Therapists were nine women and five men who were trained in CT for MDD and supervised by experts in the field (Robert DeRubeis at the University of Pennsylvania, Steven Hollon at Vanderbilt University, and David Clark and Paula Young at Rush-Presbyterian St. Luke’s Medical Center). Most therapists had low to moderate levels of experience conducting CT, with only one therapist having a very high level of experience.

Anti-Depression Medication Treatment

The ADM regiment was delivered in a flexible manner designed to represent pharmacotherapy as it would be delivered in a psychiatric specialty clinic. The general strategy, barring clinician judgment that it was not appropriate in an individual case (e.g., history of non-response to a particular ADM), was to have patients begin with 25-50 mg of sertraline (SSRI) and incremented in 50 mg steps at least every other week until they were taking a therapeutic dose at weeks 6-8. If patients did not respond after eight weeks, pharmocotherapists had an option of augmentation (e.g., with lithium) or to change to a different drug class (i.e., heterocyclics, tricyclics, monoamine oxidase inhibitors).

Session Selection

The first three CT sessions for each patient were selected. Data was omitted when recordings were not available or had been damaged; 467 sessions were successfully collected and the distribution was approximately equal across sessions (for session 1, n = 160; for session 2, n = 157; for session 3, n = 150).
Measures

Outcome: Depressive Symptoms

Depressive symptoms were evaluated with two separate measurements. The Beck Depression Inventory-II (BDI-II; Beck, Steer & Brown, 1996) is a self-report assessment that was used as the primary measure of the severity of session-to-session depression symptoms. The BDI-II is a 21-item inventory which captures affective, cognitive and somatic symptoms of depression. It is a widely used measure which has been found to be a reliable and well-validated measure of depressive symptom severity. The BDI-II was collected at the beginning of each treatment session. The modified 17-item HRSD (Williams, 1988) is a clinician administered semi-structured interview that captures somatic, affective and cognitive symptoms of depression, and also allows the interviewer to code both verbal and behavioral responses. For the purposes of defining response in the trial, the HRSD was the primary symptom outcome measure used. Therefore the slope of HRSD scores after session 3 and through week 16 was calculated and used as the outcome variable for the 16 week analyses.

Therapeutic Alliance

The alliance was assessed using the abbreviated version (Travey & Kokotovic, 1989) of the Working Alliance Inventory (WAI; Horvath & Greenberg, 1989). The WAI was developed as a transtheoretical measure of alliance and is one of the most commonly used measures of alliance to date (Martin et al., 2000). The scale is based on Bordin’s (1979) definition of the working alliance as comprised of three crucial components: Goals, which reflects the agreement between therapist and patient as to the overarching goals of treatment; Tasks, which reflects the agreement on the appropriate tasks to focus
on in order to achieve the therapeutic goals; and Bond, the quality of the affective relationship between the therapist and the patient.

The shortened form of the WAI consists of 12 items and was developed through a confirmatory factor analysis on the original 36-item scale (Horvath & Greenberg, 1989) by Tracey and Kokotovic (1989), as rated by patients and therapists. The results of the factor analysis suggested that a two-level hierarchical model achieved the best fit of the multiple models examined, presenting a measure of general alliance that can be broken down into three subscales. Keeping in line with the idea of three unique factors of Goal, Task and Bond, they took the four highest-loading items to form a shorter, more user-friendly version of the WAI. Testing the factor structure of the new shortened scale, the hierarchical model yielded acceptable fit.

The current study utilized the observer version of the WAI, which was found to have high inter-rater agreement (Tichenor & Hill, 1989). Observers have often been used to assess the alliance (e.g., DeRubeis & Feeley, 1990; Raue et al., 1994; Castonguay et al., 1996, Krupnick et al., 1996; Feeley et al., 1999); independent raters allow researchers to provide instructions and training to ensure that ratings are made according to a specific definition of the alliance. This ensures that ratings will not merely represent the overall impression of a patient – which may be more influenced by their own personal affect or social desirability – or of a therapist – which may be influenced by the degree of progress made in session according to their therapeutic orientation (Horvath & Beti, 2002). As mentioned above, Horvath and Symonds (1991) found that independent observer ratings of alliance were more predictive of outcome than therapist and patient ratings.
Coders

A total of 50 advanced undergraduate students from The Ohio State University served as independent raters. They all attended two two-hour seminars that provided them with background information on central elements of the study, including information on depression, CT and a review of the rating manual. Raters completed a training sequence comprised of watching and coding multiple CT sessions per week over a nine week period. During this time, raters also attended weekly meetings to review and discuss their ratings, and received instructional feedback as needed. This process culminated in approximately 50 hours of training over the course of 10 weeks.

Procedure

All sessions were reviewed by four independent raters who were blind to each others’ ratings and outcome. Sessions to be rated by each rater were quasi-randomly derived, with the restriction that each coder watched only one session per patient, to prevent any potential bias that may be related to knowledge of a patient’s previous therapy sessions. The order in which each rater completed ratings of his or her assigned sessions was randomized. Meetings were conducted periodically to guard against rater drift across time.
CHAPTER 3

RESULTS

Inter-Rater Reliability of Ratings

Random effects intraclass correlation coefficients (ICCs) were calculated in order to assess the inter-rater reliability of the alliance ratings. ICCs were corrected for the number of coders per session in order to derive a measure of the reliability of the average of four raters’ judgments. ICCs were: .74 for the overall alliance, .64 for the Goal subscale, .71 for the Task subscale, and .72 for the Bond subscale. As the planned long-term analyses use an average of all ratings made for each patient (across sessions), ICCs were also calculated for the average of all ratings made for each patient. These ICCs were corrected for the number of ratings made per patient (with the correction based on the harmonic mean of the number of ratings made per patient, 10.61). These ICCs estimating the reliability of the average of the ratings of each patient were: .88 for the overall alliance, .82 for the Goal subscale, .86 for the Task subscale, and .87 for the Bond subscale.

Depressive Symptoms in the Sample

At the intake assessment, the average score on the Hamilton Depression Rating Scale (HRSD; Hamilton, 1960) was 21.8 (SD = 4.0) and the mean BDI-II score was 32.0 (SD = 10.0). Means and standard deviations for BDI-II scores at session 1 through
session 4 are provided in Table 2. As the table shows, BDI-II scores appeared to decrease over time. There was evidence of significant change in BDI-II scores from session 1 to session 4 ($t(112) = 6.42, p = .002, d = 1.2$). Given the focus of planned analyses on symptom change from one session to the next, a series of paired $t$-tests were also used to examine change over each between-session interval. As shown in Figure 1, BDI scores significantly decreased in two of the three between-session intervals examined. While scores decreased significantly from session 1 to session 2 and from session 2 to session 3, scores did not differ significantly between session 3 and session 4.

*Alliance Ratings in the Sample*

The average overall and subscale alliance scores were highly correlated (see Table 3). The distribution of the alliance scales was found to be negatively skewed. A transformation of raising the total score and subscale scores to the power of three resulted in distributions with less skew. Although these transformations resulted in distributions which were somewhat leptokurtic, the underestimates of variance associated with positive kurtosis are not substantial, except in small samples (Tabachnick & Fidell, 2001).

*Primary Analyses*

Primary analyses were designed to address three questions regarding the role of the alliance in CT when combined with pharmacotherapy: (1) Does the alliance predict subsequent session-to-session symptom change over the first three sessions of CT? (2) Does the alliance predict risk of drop out over the course of acute treatment (which could last up to one year)? and (3) Is the alliance predictive of the slope of symptom change following the first three sessions through the first 16 weeks of treatment?
Analysis of session-to-session symptom change

To examine predictors of early session-to-session symptom change, the analytic approach was repeated measures regression using SAS Proc Mixed with no random effects specified (Littell et al., 2006). In these models, the relationship between the alliance in a given session and the BDI-II score at the subsequent session was assessed for all patients, controlling for BDI-II score at the current session. The dependent variable was the BDI-II score at each subsequent session (i.e., sessions 2 through 5) and the BDI-II score from the previous session (i.e., sessions 1 through 4) was entered as a covariate in the model. In other words, session 1 BDI-II would serve as a covariate in predicting the session 2 BDI-II, session 2 BDI-II would then serve as a covariate in predicting session 3 BDI-II, etc. Each process measure (assessed at sessions 1, 2, and 3) was entered into the model to serve as a predictor of subsequent session BDI-II, controlling for the current session BDI-II. Two other potential covariates were also considered for these models: therapist and site. Therapist was not found to be a significant predictor of session-to-session change ($p > .2$), but site was found to be a significant predictor ($F(157) = 5.73, p = .004$), thus site was also included in the model as a covariate in all session-to-session analyses.

As shown in Table 4, overall alliance scores just failed to reach significance, but the non-significant trend suggests that high levels of overall alliance may predict a subsequent decrease in depressive symptoms at the next session. The Task subscale was a significant predictor of session-to-session symptom change; the agreement on the tasks required to achieve therapeutic goals predicted greater symptom improvement at the next session. The effect sizes for both of these findings were small in magnitude. Although the
effects were in the expected direction, neither the Goal nor the Bond subscales achieved statistical significance (or even non-significant trends) in predicting short-term improvement. When all three subscales were added into the statistical model together, the Task subscale became a nearly-significant trend ($p = .053$), suggesting that it may be contributing a portion of variance above and beyond the common variance shared between the three related factors. 

Analysis of drop-out

Patients were coded as dropping out of treatment if they discontinued participation in the study prior to meeting remission criteria and prior to the end of one year of treatment. Using these criteria, a total of 14 patients (7.9%) in the current sample were identified as dropouts; on average, patients who dropped out remained in the study for 107.9 days (SD = 61.2 as compared to the average length of 94.2 days for those patients who met response criteria within the first year (see Figure 2 for more details)). To examine the alliance as a predictor of risk for drop-out, alliance scores were averaged across sessions. These average scores for each patient were then standardized ($M = 0, SD = 1$). Standardized overall alliance scores were examined as predictors of risk of drop-out using logistic regression (using SAS PROC Logistic). Intake BDI-II scores were included as a covariate to guard against the possibility that any relationship detected between alliance and drop-out would not be due only to patients with different levels of initial severity forming either a more or less positive alliance with their therapists. Intake BDI-II, however, was not found to predict drop-out ($OR = .99, \chi^2 = .15, p = .69$).
As Table 5 shows, the neither the overall alliance score nor any of the three alliance subscales were found to be significant predictors of risk of dropout in the first year of treatment.

*Analysis of longer-term symptom change*

Longitudinal random coefficients analysis (conducted using PROC Mixed) was utilized to determine if average alliance scores at the first three sessions would predict the slope of subsequent symptom change following the first three sessions through the first 16 weeks of treatment. Although the intention of the clinical trial was to see every patient through to remission, allowing the possibility of over one year of acute treatment, the time frame was shortened for these analyses for three reasons. First, it is unlikely that the effects of the alliance would extend for an entire year. Second, if the time period was extended to be as long as one year, any relationships detected might have been due to a multitude of intermediate therapy techniques or life experiences which might influence the course of treatment over that long period. Finally, 16 weeks has been a common length of acute treatment in previous studies of CT. In these analyses, the intercept and slope of HRSD scores were modeled as random effects. The primary effect of interest is the interaction of an alliance measure and time. A significant negative effect would indicate that alliance scores are associated with a steeper (i.e., more rapid) subsequent decline in depressive symptoms. An interaction of intake HRSD and time was entered as a covariate to control for any relationship between initial symptom severity and the slope of symptom change following session 3 through week 16.

Table 6 presents the results of the longitudinal random coefficients analyses predicting the slope of HRSD scores following the first three sessions through the first 16
weeks of treatment. The overall alliance score as well as the Task and Goal subscales were all significantly predictive of slope with small effect sizes. The negative direction of the $t$-scores indicates that higher ratings of the overall alliance, Task and Goal are associated with a steeper decline of HRSD scores; in other words, as these alliance scores increase, the rate of symptom improvement is greater. As with the drop-out analyses, when all three subscales were included into the model, no significant effects were found for any of the individual subscales, implicating the substantial common variance between the subscales as driving the findings.

**Secondary Analyses**

*Symptoms predicting session-to-session changes in alliance*

In addition to predicting subsequent symptom change from measures of the alliance, analyses were also run to test if the reverse was true. That is, the session-to-session model detailed above was reversed to explore whether current symptom levels (i.e., BDI-II scores) would predict alliance scores at the next session, controlling for the current alliance rating. As can be seen in Table 7, BDI-II does significantly predict the overall alliance and all three subscales at the subsequent session. In fact, the effect sizes for these predictions, while still in the small range, are as large as (and slightly numerically larger than) those that were found for the alliance predicting subsequent symptom change. This suggests that tests of the alliance-outcome relationship in which the temporal precedence of the alliance to outcome is not established may be inflated by virtue of change in outcome leading to improvements in alliance scores.
CHAPTER 4

DISCUSSION

In this study, observer ratings of the alliance in CT were found to predict session-to-session symptom improvements over the first three sessions (at the level of a non-significant trend) and symptom improvement over a 16-week period. Consistent with previous research, these effects were small in magnitude. The alliance ratings from early sessions, however, were not found to predict the probability of patients prematurely dropping out of treatment. Secondary analyses also revealed that early symptom change predicted the strength of the alliance, supporting the argument for designing studies of the alliance-outcome relationships so that subsequent change can be measured distinctly (as there is otherwise a danger of misinterpreting analyses due to reverse causality).

While two meta-analyses (Horvath & Symonds, 1991; Martin et al., 2000) concluded that a consistent positive relationship has been found between the alliance and outcome, these results are difficult to interpret because they include a combination of treatments, populations, and methodologies, making generalization complicated. Further, it is possible that important differences do exist between the studies included in the meta-analyses, but they may not have been detected because of insufficient numbers within various types of treatment studies. Even within studies that specifically examined cognitively oriented psychotherapies, significant results have been somewhat inconsistent
(e.g., Feeley et al., 1999; Klein et al., 2003). Again, methodological differences and shortcomings may have contributed in part to these inconsistencies; some studies that found a positive alliance-outcome association (e.g., Castanguay et al., 1996) did not take into consideration how symptom improvement prior to the alliance assessment may inflate the alliance–outcome relationship. Other studies that did focus their investigation on subsequent symptom change failed to find a significant relationship between alliance and outcome (e.g., DeRubeis & Feeley, 1990), but most of these studies consisted of very small sample sizes and may have been underpowered to detect the association of interest. Therefore, another specific intention of the present study was to address some of the methodological concerns from past research, including not only the employment of a suitable method of measurement, but also the establishment of temporal precedence of these measures in a large, diagnostically homogenous sample. Three primary hypotheses were examined, and each will be discussed in turn below.

The first research question was whether the alliance was predictive of early session-to-session symptom improvement. It was proposed that overall alliance would be predictive of early symptom improvement. In partial support of this hypothesis, a non-significant trend was found for the overall alliance ($r = .15$), but a significant main effect was found for the Task subscale ($r = .18$) in predicting subsequent symptom change. These effect sizes are slightly smaller than the effect size estimates of meta-analytic reviews of the alliance-outcome association, which ranged from $r = .22$ to $r = .26$ (Horvath & Symonds, 1991; Martin et al., 2000). However, they are quite comparable to the mean effect size of five studies specifically examining the alliance as a predictor of subsequent symptom change in people with MDD (see Table 1 for details). The weighted
mean effect size of the alliance-outcome relation in those studies is an $r$ of .12 (with a range from $r = -.27$ to $r = .15$). Only one of those studies specifically examines the combination of CT+ADM for MDD – in that study, Feeley and colleagues (1999) examined 25 patients who were either participating in CT alone or CT in combination with ADM. They found a negative effect size of $r = -.27$ for the alliance-outcome association. The study reported by Feeley and colleagues, however, was largely underpowered (to detect an effect of .15) and was not sufficiently large to allow for a meaningful examination of the alliance-outcome association specifically among patients participating in both CT and ADM.

The only other study that analyzed short-term outcome failed to find that the alliance, or any of the WAI subscales (Strunk et al., 2009), were associated with session to session change. However, the effects sizes in that study (overall alliance $r = .15$, Task $r = .17$, Goal $r = .24$, Bond $r = .06$) were similar to those obtained in the current study (overall alliance $r = .15$, Task $r = .18$, Goal $r = .11$, Bond $r = .10$). When considering the long-term analyses, Klein and colleagues (2003) found a significant association between alliance and outcome with a small effect size ($r = .14$), whereas Strunk and colleagues (2009) failed to obtain significance with a slightly smaller effect ($r = .10$). It is possible that these differences in results were due to the variations in treatment types. The treatment in the study by Klein and colleagues was Cognitive-Behavioral Analysis System of Psychotherapy (CBASP; McCullough, 2000) which, while based on cognitive strategies, places a stronger emphasis on developing the therapeutic relationship. Therefore, it is not surprising that a significant relationship was found between the alliance and outcome within CBASP. Strunk and colleagues (2009), like the current
A study, focused on CT for people with MDD, but there were two main differences from the current sample: (1) the patients were all suffering from severe depression and (2) CT was provided as a monotherapy. The symptom severity may have played a role in the negative findings by Strunk and colleagues, whereas the current study showed a greater variation in symptom levels and established a significant association between the Task subscale and short-term symptom change. It is possible, then, that the strength of the alliance, namely a collaborative agreement on Tasks, early in treatment may be predictive of outcome only when symptoms are mild to moderate in severity. However, additional exploratory analyses failed to reveal any evidence of initial severity serving as a moderator of the alliance-outcome relationship in the present study. As the CT in the current study is in combination with pharmacotherapy, it could be assumed that the psychosocial aspect may represent a smaller portion of all the therapeutic efforts undertaken. Therefore, the association between the alliance and outcome is surprisingly consistent with the literature on CT alone, which suggests that this relationship is either quite robust and can be seen in various treatments, or perhaps the alliance should be considered as a process variable largely dependent on patient characteristics that presents itself across different treatment approaches.

These findings support the notion that it is important to establish a collaborative relationship early in CT, a treatment in which therapists are encouraged to work with patients to reach agreement on the tasks patients can actively undertake to impact their negative mood states. As the general focus of CT is to be a short-term treatment aimed at teaching patients skills to directly alter their maladaptive thoughts, feelings and behaviors, the key role of the therapist, particularly in the first few sessions, is to create a
working relationship centered around the active participation required on part of the client. The current data suggests that, in order to influence early treatment gains, the focus should be less on emphasizing the affective relationship (Bond) or the abstract ideals for the future (Goal), but instead on instilling confidence in the treatment techniques presented in CT and how one can influence change in the here and now (Task), such as reevaluating negative thoughts and increasing pleasurable activities. While tempting to draw strong inferences across subscales, however, it is important to keep in mind that the effect sizes were not dramatically different and subscales were all highly correlated with one another.

The second hypothesis investigated whether the alliance was predictive of drop-out over the acute treatment phase. Neither the overall alliance nor any of the three subscales were found to be significant predictors of drop-out. Despite these negative findings, however, it is important to keep in mind that only a total of 14 patients dropped out of treatment in the interval examined. Therefore, the analyses examining predictors of drop-out may have been underpowered to detect an effect that would be of interest.

The third and final hypothesis explored whether the alliance early in treatment could predict if patients were likely to achieve symptom improvement in the first 16 weeks of treatment. The overall alliance, as well as both the Task and Goal subscales, was found to be predictive of the slope of symptom change following the first three sessions through the 16th week of treatment. This points to the importance of establishing a focus on treatment goals and active participation in tasks early in treatment, as patients are more likely to experience a steeper decline in depressive symptoms than patients that did not establish a strong working relationship in the first several sessions.
It is important to keep in mind that these findings are not in the context of CT as a monotherapy, but instead as a combination of CT with concurrent ADM use. The introduction of ADM is also accompanied by the introduction of other potential sources of variance, including possible changes in patient expectations and an additional potentially therapeutic relationship. Conclusions, therefore, must be drawn with caution in regard to the role of alliance within CT for MDD. The effects of simultaneous pharmacotherapy could potentially impact the findings of the alliance as a process variable. On one hand, ADM usage could increase early alliance because it might affect symptom reduction, which causes people to feel better and perhaps more easily form a working relationship. On the other hand, it is also possible that CT patients would tend to engage less fully in CT when they are concurrently participating in ADM. In addition, the alliance with one’s therapist in CT might be compromised if patients attribute symptom improvements to the ADM rather than the CT (perhaps missing opportunities for the patient to gain experiential evidence of the benefits of CT strategies). However, while there are still some remaining questions in regards to the alliance-outcome relationship in psychotherapy with adjunctive ADM, the combination of treatments is regularly seen in naturalistic settings and thus the results of this study are taking steps to increase generalizability to the population of people with MDD.

Limitations

While a major aim of the current study was to address a number of methodological issues in the literature on process research, a number of limitations remain to be addressed. First, as the study consisted entirely of patients in the combination treatment of CT+ADM, it makes it difficult to directly compare and relate
the findings to CT as a monotherapy for MDD. Means could be compared across similar studies and estimates of the alliance-outcome relationship could be obtained, but it would not be as precise as if the data had been collected simultaneously in the same research study. Even if separate treatment studies had the same intentions, the data would likely end up being different as the samples are not the same, in number or in character, and the raters may be differentially biased, making it inappropriate to compare across studies. The two studies measuring only subsequent symptom change that included both a combination treatment group and a psychotherapy alone group have methodological differences that also make direct comparisons difficult: Feeley and colleagues (1999) did not have sufficient enough power to separate the CT alone from the CT+ADM conditions in their analyses and Klein and colleagues (2003) utilized a form of cognitive-behavioral therapy that, while similar in some respects to CT, emphasizes an interpersonal focus and thereby potentially boosting the strength of both the therapeutic alliance and the alliance-outcome association. Thus, a clear goal for future research on the role of the alliance in treatment is to employ a study that directly compares CT alone to CT+ADM in treatment of MDD, with a randomization to conditions and a large enough sample size to achieve statistical power to confidently address significance and utilizing the same raters of the alliance across treatment groups, blind to conditions.

Another potential limitation of the current study is that the independent raters of the alliance were all undergraduate students. While these students did not have graduate level training in psychotherapy, they were also unlikely to have strong beliefs about any specific theoretical orientation. These students, however, were all advanced undergraduates who expressed interest in pursuing higher education in psychology, and
there is no compelling reason to believe that advanced training is required to rate the alliance. Relative to graduate students or professionals, the undergraduates may have come in to the study with fewer preconceived notions regarding what therapy “should” look like, as they were not tied to one therapeutic orientation. For example, if the raters had an opinion regarding who was likely to benefit from treatment according to their perception of the treatment-specific quality of the session that they watched, they could bias their ratings (e.g., artificially inflate the ratings of “good” sessions and deflate ratings of “bad” sessions) so as to increase the alliance-outcome association. Any possible shortcomings due to the current raters’ lack of experience, therefore, were attenuated by extensive training and repeated exposure to CT sessions. And, while the reliability of a single rater was not high (overall alliance ICC = .42), this effect was minimized by increasing the reliability with multiple ratings (4) per session (overall alliance ICC = .74).

A third limitation related to the choice of sessions to rate. While there was a clear rationale for choosing the first three sessions, it is possible that there is meaningful variance that is not being captured by keeping the selection so limited, such as the evolution of the therapeutic alliance over time, which may include potential rifts in the relationship. It is also possible that the first few sessions are categorically different from later sessions in that the relationship is still in its very early development, which does not represent how the alliance persisted over the course of treatment. There is evidence, for example, that patient symptom improvement does predict an increase in the therapeutic relationship over time (the current study; Strunk et al., 2009; DeRubeis & Feeley, 1990). The intention, however, was to capture the early alliance as it formed in order to identify any information that may be useful in establishing the initial working relationship and
encouraging early therapeutic gains. Another benefit of selecting these first three sessions is that each patient was started on the same course of ADM, therefore minimizing the amount of variance attributed to differences in medication regimes at this point, as there would likely be in later sessions in this trial. Future studies need to have a similar, clear rationale for the sessions that are selected to be rated so that analyses and comparisons between studies make sense, and it would be interesting to compare our findings with a comparable study focusing on a collection of late sessions.

A fourth limitation is that third variables could be playing an important role in the causality of these findings. While there are limitless possibilities of confounding relationships that make it difficult to account for this source of variance, further research can be done with certain pretreatment characteristics that may be of interest, such as depression chronicity, the presence of comorbid diagnoses, and other factors that may complicate or contribute to the strength of the developing alliance. Nonetheless, exploratory analyses of four variables (i.e., age of onset, initial severity, number of previous depressive episodes, and the presence of an Axis II disorder) were examined and there was no evidence that they accounted for the associations of interest in the current paper.

A final limitation concerns the measurement of the alliance itself. There are various conceptualizations of the therapeutic alliance, which are captured in a variety of measurement techniques, which may differ significantly from the operationalization of alliance presented in this paper. Recently, for example, there has been interest in recognizing the importance of alliance ruptures, which emphasizes that attending to and successfully resolving negative interpersonal problems between the therapist and patients
in treatment could be predictive of outcome (Ruiz-Cordell & Safran, 2007). The emphasis, then, is not on the strength of the therapeutic relationship, but instead on the ability to openly resolve potential conflict. This opens the possibility that the alliance is a multi-faceted concept that could be defined and studied in a variety of ways.

**Conclusions and Future Implications**

In a sample of 176 patients with MDD undergoing a course of combined CT+ADM treatment, overall alliance was nearly predictive (and the Task subscale was a significant predictor) of early treatment gains, which is when the majority of symptom improvement is seen. The alliance and two subscales (Task, Goal) were also found to predict the rate of improvement over a 16 week period in treatment. The alliance was not found to be a significant predictor of treatment retention. The magnitudes of the findings involving the alliance predicting symptom change are quite small, but are consistent with the literature. Thus, though it will be important to identify other predictors, these findings suggest efforts to develop and maintain a positive therapeutic relationship may be important for maximizing symptom change.
REFERENCES


FOOTNOTES

1. Another small study conducted by Gaston, Marmar, Gallagher, and Thompson (1991) also failed to find an alliance-outcome association in a sample of older adults with MDD receiving either CT (n = 17), behavioral therapy (BT; n = 16) or brief psychodynamic therapy (B-PT; n = 21). At first glance, this study appeared to measure subsequent change in depression symptoms, but the authors did not control for the current symptom level, so these results were not included in the group of studies with temporal considerations.

2. All significant results non-significant trends found with the transformed data remained significant when run with the raw data.

3. Because the amount of change in BDI scores varied across sessions, I also examined whether there was any evidence of the alliance-outcome relationship varying across sessions. There was no evidence of an interaction between session and alliance in the session-to-session model, suggesting that the alliance-outcome relationship did not vary significantly across the sessions examined.
APPENDIX A

TABLES
Table 1. Studies of Alliance-Outcome Associations with Temporal Considerations that Measure Only Subsequent Symptom Change

<table>
<thead>
<tr>
<th>Authors</th>
<th>Measure</th>
<th>Sample</th>
<th>n</th>
<th>Treatments</th>
<th>ES ($r$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>DeRubeis and Feeley, 1990</td>
<td>HAq</td>
<td>MDD</td>
<td>25</td>
<td>CT</td>
<td>.10</td>
</tr>
<tr>
<td>Feeley, DeRubeis, and Gelfand, 1999</td>
<td>HAq</td>
<td>MDD</td>
<td>25</td>
<td>CT, CT+ADM</td>
<td>-.27</td>
</tr>
<tr>
<td>Barber et al., 1999</td>
<td>CALPAS</td>
<td>Cocaine Dependent</td>
<td>252</td>
<td>SE, CT, B-PD, 12-Step</td>
<td>.01</td>
</tr>
<tr>
<td>Barber et al., 2000</td>
<td>CALPAS</td>
<td>GAD, MDD, OCPD, APD</td>
<td>86</td>
<td>SE</td>
<td>.30*</td>
</tr>
<tr>
<td>Strunk, Brotman, &amp; DeRubeis, 2009</td>
<td>WAI</td>
<td>MDD (Severe)</td>
<td>60</td>
<td>CT</td>
<td>16 week: .10 early sess: .15</td>
</tr>
<tr>
<td>Zuroff &amp; Blatt, 2006</td>
<td>VTAS</td>
<td>MDD</td>
<td>161</td>
<td>CT, IPT, ADM+CM, PLA+CM</td>
<td>.15*</td>
</tr>
<tr>
<td>Klein et al., 2003</td>
<td>WAI</td>
<td>MDD (Chronic)</td>
<td>367</td>
<td>CBASP, CBASP+ADM</td>
<td>.14*</td>
</tr>
</tbody>
</table>

Note: Measures: HAq = Penn Helping Alliance Questionnaire, CALPAS = California Psychotherapy Alliance Scale, WAI = Working Alliance Inventory, VTAS = Vanderbilt Therapeutic Alliance Scale; Sample: MDD = Major Depressive Disorder, GAD = Generalized Anxiety Disorder, OCPD = Obsessive Compulsive Personality Disorder, APD = Avoidant Personality Disorder; Treatments: CT = Cognitive Therapy, B-PD = Brief Psychodynamic Therapy, SE = Supportive Expressive Therapy, ADM = Anti-Depressant Medication, CBASP = Cognitive-Behavioral Analysis System of Psychotherapy, IPT = Interpersonal Psychotherapy, CM = Clinical Management, PLA = Placebo
Table 2. Average BDI-II Scores Across Sessions

<table>
<thead>
<tr>
<th></th>
<th>N</th>
<th>Mean</th>
<th>Standard Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Session 1</td>
<td>134</td>
<td>27.3</td>
<td>11.1</td>
</tr>
<tr>
<td>Session 2</td>
<td>135</td>
<td>23.4</td>
<td>11.4</td>
</tr>
<tr>
<td>Session 3</td>
<td>127</td>
<td>21.8</td>
<td>11.0</td>
</tr>
<tr>
<td>Session 4</td>
<td>150</td>
<td>21.8</td>
<td>12.3</td>
</tr>
</tbody>
</table>
Table 3. Correlations Between the Mean Alliance for Individual Patients

<table>
<thead>
<tr>
<th></th>
<th>Overall Alliance</th>
<th>Task</th>
<th>Goal</th>
<th>Bond</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overall Alliance</td>
<td>--</td>
<td>.97</td>
<td>.91</td>
<td>.94</td>
</tr>
<tr>
<td>Task</td>
<td>.97</td>
<td>--</td>
<td>.83</td>
<td>.89</td>
</tr>
<tr>
<td>Goal</td>
<td>.91</td>
<td>.83</td>
<td>--</td>
<td>.77</td>
</tr>
<tr>
<td>Bond</td>
<td>.94</td>
<td>.89</td>
<td>.77</td>
<td>--</td>
</tr>
<tr>
<td>M</td>
<td>49.08</td>
<td>15.26</td>
<td>17.76</td>
<td>16.06</td>
</tr>
<tr>
<td>SD</td>
<td>6.35</td>
<td>2.56</td>
<td>2.14</td>
<td>2.03</td>
</tr>
</tbody>
</table>

Note: All items are rated on a scale from 0 (not at all) to 6 (extensively); Overall alliance scores have a potential range of 0-72 and each of the subscales could range from 0-24.
Table 4. Association between Alliance and Symptom Improvement in the Following Session, Covarying Site and Current BDI-II score

<table>
<thead>
<tr>
<th></th>
<th>$r$</th>
<th>$t$</th>
<th>$p$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overall Alliance</td>
<td>.15</td>
<td>-1.89</td>
<td>.061</td>
</tr>
<tr>
<td>Task</td>
<td>.18</td>
<td>-2.22</td>
<td>.028</td>
</tr>
<tr>
<td>Goal</td>
<td>.11</td>
<td>-1.42</td>
<td>.157</td>
</tr>
<tr>
<td>Bond</td>
<td>.10</td>
<td>-1.22</td>
<td>.224</td>
</tr>
</tbody>
</table>

Note: The negative $t$-scores indicate that an increase in Alliance is associated with a decrease in symptom scores at the following sessions. The $r$ scores indicate the strength of association between the alliance at session $n$ and the symptom level at session $n + 1$, controlling for symptoms at session $n$. All $df = 152$. 

58
Table 5. Average Alliance as a Predictor of Drop-out, Covarying Intake BDI-II Scores

<table>
<thead>
<tr>
<th></th>
<th>Odds Ratio</th>
<th>Wald Chi</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overall Alliance</td>
<td>.801</td>
<td>.626</td>
<td>.429</td>
</tr>
<tr>
<td>Task</td>
<td>.746</td>
<td>1.01</td>
<td>.315</td>
</tr>
<tr>
<td>Goal</td>
<td>.891</td>
<td>.167</td>
<td>.683</td>
</tr>
<tr>
<td>Bond</td>
<td>.802</td>
<td>.614</td>
<td>.433</td>
</tr>
</tbody>
</table>

Note: The odds ratios less than 1 indicate a decreased likelihood of dropping out of the study during the acute treatment phase
Table 6. Alliance as a Predictor of the Slope of HRSD scores following Session 3 through Week 16, Covarying Intake HRSD

<table>
<thead>
<tr>
<th></th>
<th>r</th>
<th>t</th>
<th>df</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overall Alliance</td>
<td>.20</td>
<td>-2.28</td>
<td>126</td>
<td>.002</td>
</tr>
<tr>
<td>Task</td>
<td>.23</td>
<td>-2.69</td>
<td>128</td>
<td>.008</td>
</tr>
<tr>
<td>Goal</td>
<td>.20</td>
<td>-2.20</td>
<td>118</td>
<td>.03</td>
</tr>
<tr>
<td>Bond</td>
<td>.12</td>
<td>-1.43</td>
<td>129</td>
<td>.16</td>
</tr>
</tbody>
</table>

Note: The negative t-scores indicate that an increase in Alliance is associated with a steeper slope (i.e., a more rapid decrease in symptoms)
Table 7. Association between Symptom Level and Alliance in the Following Session, Covarying Site and Current BDI-II score

<table>
<thead>
<tr>
<th></th>
<th>r</th>
<th>t</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overall Alliance</td>
<td>.21</td>
<td>-2.46</td>
<td>.015</td>
</tr>
<tr>
<td>Task</td>
<td>.22</td>
<td>-2.57</td>
<td>.011</td>
</tr>
<tr>
<td>Goal</td>
<td>.20</td>
<td>-2.37</td>
<td>.019</td>
</tr>
<tr>
<td>Bond</td>
<td>.17</td>
<td>-2.03</td>
<td>.044</td>
</tr>
</tbody>
</table>

Note: The negative t-scores indicate that a decrease in symptom scores is associated with an increase in Alliance at the following sessions. The r scores indicate the strength of association between the symptom level at session n and the Alliance at session n+1, controlling for Alliance at session n. All df = 136.
APPENDIX B

FIGURES
Figure 1. Change scores in BDI-II over Early Between-Session Intervals

Note: White data points indicate a significant change in scores between sessions ($p < .05$); $s1 =$ session 1; $s2 =$ session 2; $s3 =$ session 3; $s4 =$ session 4
<table>
<thead>
<tr>
<th></th>
<th>Completers</th>
<th>Dropouts</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>N</strong></td>
<td>50</td>
<td>14</td>
</tr>
<tr>
<td>Median (days)</td>
<td>84.5</td>
<td>115.0</td>
</tr>
</tbody>
</table>

**Note.** The box represents the interquartile range (from the 25th to the 75th percentile), the line across the box represents the median and the lines extending from the box represent the highest and lowest values.

Figure 2. Length of Time in Treatment