Attitudes Toward Cognitive and Behavioral Interventions: Prediction of Preference and Outcomes in the Treatment of Major Depression

DISSERTATION

Presented in Partial Fulfillment of the Requirements for the Degree Doctor of Philosophy in the Graduate School of The Ohio State University

By

Andrew Astley Cooper

Graduate Program in Psychology

The Ohio State University

2013

Dissertation Committee:

Daniel R. Strunk PhD, Advisor

Jennifer S. Cheavens PhD

Mary Fristad PhD
Copyrighted by
Andrew A. Cooper
2013
Abstract

In their influential dismantling study, Jacobson and colleagues (1996) established a precedent for comparing the cognitive and behavioral interventions that comprise CBT for depression. Follow-up analyses of patients from that study suggested that certain patient characteristics might be associated with attitudes toward, and response to, these interventions. In particular, agreement with the treatment rationale and complexity of beliefs about the cause of one’s depression were identified as potential predictors, with subsequent research also suggesting a role for patient preferences. The present study investigates attitudes toward cognitive and behavioral interventions for depression, as well as complex beliefs about the cause of one’s depression, as predictors of preference and treatment outcomes.

Dysphoric students (n=193) and depressed outpatients (n=38) completed ratings of written rationales and videotaped vignettes describing the two interventions. Patients also participated in a small trial comparing these interventions as separate treatments, with symptom change and dropout assessed across the first eight weeks of therapy. All participants completed ratings of intervention credibility (using the intervention credibility scale, ICS) and personal reactions to the treatment rationales (PRR) for both interventions. Participants also rated their preference between the interventions using a continuous scale. Beliefs about the cause of their depression were also assessed using the
Reasons for Depression inventory (RFD), with higher scores reflecting a “reason-giving” tendency expected to predict less favorable reactions to the behavioral intervention.

Results were generally more consistent with hypotheses in the patient sample. Both patients and students rated the interventions as equally credible. For the patients, higher RFD scores were associated with higher absolute and relative ICS scores for the cognitive intervention, while no such associations were observed with ICS ratings in the student sample. Higher RFD scores predicted preference for cognitive treatment in the patient sample. PRR ratings of the vignettes did not differ in the student sample, but patients rated the cognitive intervention higher than the behavioral treatment. Higher RFD scores predicted lower PRR ratings of the behavioral treatment in the student sample. Difference scores on the ICS and PRR were highly predictive of preference ratings, with the higher-rated vignette being strongly preferred.

Patients tended to prefer the cognitive intervention, and thus more were “mismatched” in the behavioral treatment. Depressive symptoms generally improved across eight weeks in both treatment conditions, with a 24% dropout rate. RFD scores, ICS and PRR difference scores, preference ratings and treatment match status all failed to predict symptom change, as main effects or in interactions with treatment assignment. Higher RFD scores predicted increased odds of dropout, overall and especially in the cognitive intervention. Cognitive preference predicted higher odds of dropout overall, but the interaction of preference and treatment assignment was not significantly predictive of dropout. Finally, the interaction of PRR difference scores and treatment trended toward predicting lower dropout odds. Probability of dropout decreased as PRR ratings of a
patient’s assigned treatment were lower relative to the other treatment. These results suggest a potentially meaningful role for PRR ratings and RFD scores as predictors of preference, as well as odds of dropout. These findings are considered relative to limitations of the present study, implications for patient care, and future research directions.
Dedication

For my family.
Acknowledgments

In gratitude to the friends, family, mentors and colleagues who made this possible: I have been tremendously fortunate to receive your support, your wisdom, your advice, and your presence in my life. I am extremely grateful for the feedback, direction and encouragement I have received from my advisor, Dan Strunk, and my perennial committee members, Jennifer Cheavens & Mary Fristad, during my time at OSU.

Special thanks to my labmates (and labmates-by-proxy) and friends near and far for supplying ideas, outlets and the genuinely amazing feeling of being supported 24/7. I also owe a debt of gratitude to my incredible research assistants, and the excellent B&C study team therapists for enduring the data collection process (and enduring me, in the process) to make the pilot study a success.

I cannot thank my parents and grandparents enough for supporting me and encouraging me in every conceivable way. And finally, love and gratitude to Marie, for being the best motivator and partner I could ever ask for.
Vita

2006 ................................................................. Henry Street High School

2006 ................................................................. B.Sc., Psychology (summa cum laude)

McMaster University

2007 - 2008 ......................................................... University Fellow, Department of

Psychology, The Ohio State University

2008 - 2009 ........................................................... Graduate Teaching Assistant, Department of

Psychology, The Ohio State University

2010 ................................................................. M.A., Psychology

The Ohio State University

2009 - 2011 ........................................................... Social Sciences and Humanities Research Council (SSHRC) Doctoral Fellow

2011 ................................................................. Doctoral Candidate, Department of

Psychology, The Ohio State University

2011 - 2012 ........................................................... Graduate Teaching Assistant, Department of

Psychology, The Ohio State University

2012 to present ..................................................... Clinical Psychology Pre-doctoral Internship,

Indiana University School of Medicine
Publications


Fields of Study

Major Field: Psychology
Table of Contents

Abstract ......................................................................................................................... ii

Dedication .................................................................................................................... v

Acknowledgments ....................................................................................................... vi

Vita ................................................................................................................................. vii

Publications .................................................................................................................. viii

Fields of Study ........................................................................................................... viii

Table of Contents ....................................................................................................... ix

List of Tables .............................................................................................................. xii

List of Figures ............................................................................................................ xiv

Chapter 1: Introduction ............................................................................................... 1

   CBT and the Treatment of Depression ........................................................................ 1

   Dismantling CBT by Intervention Type ........................................................................ 3

   Comparing Cognitive and Behavioral Interventions in CBT ...................................... 6

   Predictors of Response to Cognitive and Behavioral Interventions ............................ 8

   Treatment Rationale and Credibility ........................................................................... 9

   Reason-giving and Differential Response to CBT Interventions ................................ 12
Hypothesis 3: RFD and Credibility Ratings in the Prediction of Treatment Preference
Ratings

Hypothesis 4: Prediction of Treatment Outcomes from RFD scores, Credibility
Ratings, and Preference

Exploratory Analyses: Comparing Significant Predictors of Dropout

Exploratory Analyses: Post-hoc Comparison of Students and Patients

Chapter 4: Discussion

Intervention Ratings

Reasons for Depression and Reason-giving

Preference Ratings

Prediction of Outcomes from Cognitive and Behavioral Intervention Pilot Study

Limitations

Conclusions

References

Endnotes

Appendix A: Tables

Appendix B: Figures
List of Tables

Table 1. Studies of Preference/Outcome Associations in Treatment of Depression. .... 93
Table 2. Summary of Preference/Outcome Findings in Treatment of Depression .......... 94
Table 3. Means of Intervention Credibility Ratings and Personal Reactions Ratings of Vignettes & Rationales by Intervention Type and Sample. ................................. 95
Table 4. Correlations between PRR and ICS scores, by Sample. .............................. 96
Table 5. Means of ICS and PRR Difference Scores, by Sample. ................................. 97
Table 6. RFD Total scores as Predictors of Intervention ICS and PRR Ratings and Difference Scores, with Viewing Order Co-varied, in Student and Patient Samples....... 98
Table 7. RFD Total Scores and Intervention ICS and PRR Difference Scores in the prediction of Treatment Preference scores in the Student Sample. .......................... 99
Table 8. RFD Total Scores and Intervention ICS and PRR Difference Scores in the prediction of Treatment Preference scores in the Patient Sample. ......................... 100
Table 9. HRSD Data Availability, Means, and Average Time to Assessment, Overall and by Treatment Condition. ...................................................................................... 101
Table 10. Categorical Preference Ratings and Treatment Matching Status. ............... 102
Table 11. Demographic and Treatment Characteristics by Treatment Condition, with Group Comparisons. ......................................................................................... 103
Table 12. Potential Predictors of Slope of HRSD Symptom Change across the First Eight Weeks of Treatment After Covarying Medication Status ........................................ 104

Table 13. Demographic and Treatment Characteristics by Dropout Status .................. 105

Table 14. Potential Predictors of Dropout Across the First Eight Weeks of Treatment. 106
List of Figures

Figure 1. RFD Score by Treatment Assignment in the Prediction of Dropout ............... 108

Figure 2. PRR-BvC Score by Treatment Assignment in the Prediction of Dropout...... 109
Chapter 1: Introduction

CBT and the Treatment of Depression.

Cognitive behavior therapy (CBT) is among the most well studied treatments for depression (Hollon & Beck, 2004). There is evidence that CBT is efficacious in the treatment of depression, with effects comparable to anti-depressant medication during acute treatment (Strunk & DeRubeis, 2001; Hollon, Stewart & Strunk, 2006). In spite of these promising findings, response rates during acute treatment (i.e., 12 to 16 weeks) are suboptimal, with one estimate drawn from four large clinical trials indicating that just two-thirds of treatment completers and a little over half of all participants met response criteria at the end of treatment (Strunk & DeRubeis, 2001). Rate of response also varies, with some individuals responding rapidly and others not showing a robust response until later in treatment, if at all (Fennell & Teasdale, 1987). The poor rate of response underscores the importance of identifying ways to improve outcomes in CBT.

Researchers have sought to better understand how CBT achieves its effects by studying how cognitive and behavioral interventions may be differentially associated with outcomes (Strunk, Brotman & DeRubeis, 2010). These interventions have never been formally evaluated as separate treatments, but there is some evidence that they can be delivered separately, while still adhering to the general principles of a flexible CBT protocol and achieving outcomes comparable to standard CBT (Jacobson et al., 1996; Jarrett & Nelson, 1987). To directly compare these two interventions, a small pilot study was conducted at Ohio State in which patients were randomized to receive one of these
approaches for eight weeks of treatment, then re-randomized to continue or switch focus for another eight weeks. Secondary outcomes of this study will also investigate the role of patient characteristics as predictors of response to treatment, as there is reason to believe that individuals’ beliefs about their treatment assignment might also influence outcomes.

The present study concerns attitudes towards cognitive and behavioral interventions in the treatment of depression as they relate to treatment preferences and outcomes. Individual differences in reaction to the treatment rationales presented in CBT treatment are an important predictor of outcome (Fennell & Teasdale, 1987), and were predictive of short term change and outcome in a CBT dismantling study (Addis & Jacobson, 2000). In the same study, a tendency to endorse multiple reasons for depression (dubbed reason-giving) also predicted outcome, and interacted with treatment assignment to predict poorer response in a behavioral-only treatment (Addis & Jacobson, 1996). This same reason-giving tendency predicted lower credibility ratings of a behavioral treatment rationale in a non-clinical sample (Addis & Carpenter, 1999). Both credibility ratings and beliefs about the causes of depression may be associated with an individual’s willingness to enter a given treatment, which may constitute a preference if other treatment options are also presented. Treatment preferences have been shown to predict clinically relevant outcomes such as symptom change and dropout (Swift, Callahan & Vollmer, 2011). The pilot study currently being conducted provides an ideal context to evaluate preference for cognitive and behavioral interventions in this way. However, an important precursor to this work is to better understand how the concepts of credibility, reason-giving and preference are associated with one another, and how they
might be related to differential treatment outcomes in the context of cognitive and behavioral interventions.

The goal of this research was to compare cognitive and behavioral interventions in the treatment of depression. Specifically, this project focuses on ratings of credibility, personal reactions, and reason-giving tendencies, and how these may predict preference and treatment outcomes. The sections that follow establish a foundation for a comparison between cognitive and behavioral interventions in the treatment of depression. These are followed by research connecting credibility and reason-giving to studies of cognitive and behavioral treatments. Finally a potential connection between these two concepts (viz., credibility and reason-giving) and treatment preference is outlined as a prelude to a review of preference-outcome relations observed in studies involving treatments for depression.

**Dismantling CBT by Intervention Type**

The original CBT treatment manual (Beck, Rush, Emery, & Shaw 1979) emphasizes the importance of both cognitive and behaviorally-oriented approaches in treating depression. The former category includes exercises to help clients identify negative automatic thoughts (ATs), to challenge them using various techniques, and to ultimately replace them with more realistic appraisals. In later sessions, the therapist works with the client to identify the underlying negative beliefs (or schemas) that ostensibly give rise to these ATs. Behaviorally-oriented interventions such as self-monitoring, increasing pleasure and mastery experiences, and remedying social skills deficits, are also identified as important in the original CBT treatment manual, with a particular emphasis on these approaches early in treatment. Beck and others (e.g.,
Whisman, 1999) have asserted that while these behavioral techniques may not explicitly challenge negative thoughts and underlying beliefs, they are intended to advance the progress of cognitive change, the putative primary mechanism of symptom improvement per the original cognitive model. Efforts to understand how CBT achieves its effects have included investigations of how these in-session interventions might be related to symptom improvement.

In an influential study, Jacobson and his colleagues (1996) explored the relationship between the types of interventions used in CBT sessions and treatment outcomes in 150 depressed outpatients. Patients were randomly allocated to 16 weeks of treatment in one of three conditions. In the behavioral activation (BA) condition (n=56), therapists were limited to using behavioral interventions as described above (e.g., activity logging, social skills training, etc.). In the automatic thoughts (AT) condition (n=43), therapists were allowed to use all BA interventions, and also interventions focusing on assessing and modifying automatic thoughts. These interventions included using daily thought records to identify the relationship between thoughts and emotions, and challenging negative thoughts or cognitive errors. The final condition was CBT in its complete form (n=50), which allowed the therapists to freely use BA and AT interventions as noted above. This condition not only allowed therapists to address underlying schemas, but obligated them to do so for at least 8 sessions. All three therapy protocols were designed to be implemented across 20 sessions. An adapted version of the treatment rationale was created for use in the BA-only condition, in the interest of presenting a rationale that was consistent with the main focus and activities of treatment. As is typical in CBT, rationales were presented early in treatment, in this case as part of
assigned readings (the pamphlet *Coping with Depression*; Beck & Greenberg, 1974) and planned discussions in early sessions. The BA rationale emphasized the role of reinforcing, positive experiences and evaluation of behavior change, in contrast to the traditional CBT rationale that is more focused on the influence of negative automatic thoughts and underlying beliefs.

Patients were categorized as improved and recovered if they no longer met criteria for major depression and had Beck Depression Inventory (BDI; Beck, Ward, Mendelson, Mock & Erbaugh, 1961) scores lower than eight; patients who were no longer depressed but had BDI scores of eight or greater were considered improved but not recovered. The results of the study indicated no treatment (AT, BA or CBT) differences in outcome on either the BDI or Hamilton Rating Scale for Depression (HRSD; Hamilton, 1960). Similarly, there were no significant differences by treatment for improvement (62.3% for the whole sample) or recovery (51.5% for the whole sample), though dropouts were less likely to improve or recover across all groups (16.7% improved and 5.6% recovered). The same result was observed at 6-month follow-up, with all three treatments showing equivalent outcomes in terms of BDI or HRSD, recovery status, relapse and number of “well weeks” since ending acute treatment (as assessed by the Longitudinal Interval Follow-up Evaluation; Keller et al., 1987). A follow-up study (Gortner, Gollan, Jacobson & Dobson, 1998) reporting on 137 of the original participants two years after the end of acute treatment also reported no significant differences by condition.

As the magnitude of symptom improvement was comparable to prior trials of CBT, Jacobson and colleagues interpreted their results as indicative of equivalence in outcomes between all three versions of CBT delivered in the trial, both during the acute
treatment phase and at 6-month follow-up. The authors suggested that the results of their study gave rise to important questions about the necessary and sufficient elements of CBT in terms of the interventions utilized by therapists. They speculated about the advantages of BA interventions, including the potential for enhanced cognitive change via natural reinforcement (versus challenging thoughts), and the parsimony and cost efficiency afforded by interventions that might be more accessible to less experienced therapists (for a discussion see Longmore & Worrell, 2007). Subsequent publications (e.g., Jacobson, Martell & Dimidjian, 2001) led to the development of a behavioral activation protocol and treatment manual (Martell, Addis & Jacobson, 2001) based on behavioral principles adapted from CBT and other therapies. Behavioral activation therapy appears to be at least equivalent to CBT in terms of acute symptom change and reduced risk of relapse in treating depression (Ekers, Richards & Gilbody, 2008).

**Comparing Cognitive and Behavioral Interventions in CBT.**

The Jacobson et al. (1996) trial sets up an implicit comparison between cognitively- and behaviorally-oriented interventions in CBT, but does not provide a direct comparison between the two. This is because the AT and CBT conditions included both of these interventions, especially early in treatment, with no stipulation that AT and BA interventions should be delivered separately. Indeed, process ratings conducted on randomly sampled sessions from early, mid, and late treatment suggest that use of behavioral interventions was high in all three treatments, and therapists in the AT and CBT conditions appeared to use both approaches at all three time periods. Thus, though the dismantling study was highly influential in isolating behavioral interventions from
other aspects of CBT, it cannot fully address the question of the differences between cognitive and behavioral treatments.

Although a formal comparison of cognitive and behavioral interventions as separate treatments has never been reported in the literature, there is some evidence that both interventions can operate independently and achieve positive outcomes. Evidence for behavioral-only interventions comes from both the dismantling study (Jacobson et al., 1996) and from more recent studies of the full-BA treatment developed as part of that trial (for a review of these studies, see Ekers et al., 2008). CBT treatments focusing exclusively on cognitive techniques have not been as well represented in the literature. Jarrett and Nelson (1987) reported on a group CBT treatment for 37 depressed outpatients, in which therapists were restricted to using specific cognitive interventions (e.g., use of thought records to monitor automatic thoughts; challenging negative thoughts and assumptions; testing out negative beliefs via experimentation) at certain sessions. Use of these activities was associated with symptom change and other markers of improvement, including less dysfunctional thoughts, more pleasant events and improved interpersonal relations. Thus, the results obtained by Jarrett and Nelson (1987) suggest that cognitive interventions might promote change when delivered independent of behavioral interventions, a possibility not fully explored by the Jacobson et al. (1996) study.

A pilot study was conducted in order to directly compare cognitive and behavioral interventions. In the initial phase of this study, depressed outpatients were randomized to receive eight weeks of treatment with either cognitive interventions (i.e., a focus on changing automatic thoughts, utilizing techniques such as thought records and hypothesis
testing) or behavioral interventions (i.e., a focus on establishing reinforcing experiences and challenging problematic behaviors, utilizing techniques such as activity logging and social skills training). As in the dismantling study, patients in this pilot study were presented with treatment rationales that are consistent with the current focus of their treatment (e.g., a focus on challenging negative automatic thoughts in the cognitive condition). The pilot study was meant to evaluate the relative efficacy of cognitive and behavioral interventions in treating depressive symptoms, along with various other outcomes of interest (e.g., dropout). This manuscript includes a subset of the outpatient population that was recruited for the pilot study, and considers the role of patient characteristics as prescriptive predictors of treatment outcomes, a target of considerable interest for CBT researchers (Jarrett, Eaves, Grannemann & Rush, 1991; Hamilton & Dobson, 2002; Coffmann, Martell, Dimidjian, Gallop & Hollon, 2007). A number of studies of this kind have also explored the role of patient characteristics as moderators of outcome on the basis of treatment assignment (Beutler et al., 1991; Sotsky et al., 1991; Fournier et al., 2009). Two follow-up studies to the CBT dismantling trial conducted by Addis and Jacobson (1996; 2000) highlight the importance of considering individual differences as predictors of response to treatment, even when overall treatment-level differences are not detected. In the case of these studies, the differences in question occur in patient beliefs about depression and treatment.

**Predictors of Response to Cognitive and Behavioral Interventions**

In the following sections, three patient characteristics will be introduced as potential ways of understanding individual differences in response to cognitive and behavioral interventions: acceptance of the treatment rationale; beliefs about the reasons
for one’s depression (especially a tendency to endorse multiple causes); and preferences for treatment. The first two concepts are directly connected to the CBT dismantling trial by virtue of the aforementioned follow-up studies, while the third idea (preferences) is strongly suggested by the results of those studies, and more generally by research on treatments for depression. While each of these concepts has been investigated in prior research involving CBT, the focus of this project is to evaluate how these ideas are related to one another and predictive of outcomes in the context of a comparison of cognitive and behavioral interventions. The literature review that follows attempts to establish these ideas as potential predictors of clinical outcomes such as symptom change or dropout, and also as potentially informative with respect to how individuals who are experiencing some symptoms of depression might evaluate potential treatment options.

**Treatment Rationale and Credibility**

Addis and Jacobson (2000) explored the degree to which a patient’s agreement with the treatment rationale was predictive of superior outcomes in the CBT dismantling trial, inspired in part by an influential early study of CBT by Fennell and Teasdale (1987). In that study, judges’ ratings of client agreement with the CBT model were predictive of symptom improvement at post-treatment and at two post-treatment follow-up assessments. As Addis and Jacobson (2000) reported, patients in the dismantling trial were asked to respond to the question, “To what degree does the treatment you are receiving match with your ideas of what helps people in psychotherapy?”, with ratings scaled from 0 (not at all) to 100 (completely) and averaged across the first three sessions. These sessions were selected to evaluate initial reactions, and coincided with planned discussion of the patients’ reactions to reading the *Coping with Depression* pamphlet.
Ratings of acceptance of the treatment rationale were predictive of homework compliance, symptom change from intake to both early (sessions 4-6) and mid-treatment (sessions 10-12) as assessed by the BDI, and treatment outcome (based on a composite of HRSD and BDI scores). Somewhat surprisingly, in their investigation of acceptance of the treatment rationale, Addis and Jacobson (2000) elected to aggregate all three treatment groups from the dismantling study into a single sample, despite differences in both the theoretical rationales and the actual materials presented to participants in the three conditions (i.e., differences in the *Coping with Depression* pamphlet described earlier).

Addis and Jacobson (2000) used a single item to assess agreement with the treatment rationale, but there are limitations to using single items in research (Gardner, Cummings, Durham & Pierce, 1998). The extensive literature on treatment evaluation suggests a suitable and more comprehensive alternative, based on an analogous concept: treatment credibility. Rokke and colleagues (Rokke, Carter, Vehm & Veltum, 1990) developed a measure of this construct based on prior work by Borkovec and Nau (1972) which identified dimensions of treatments thought to be informative in generating expectancy for change. Their original credibility scale consisted of nine items rated on a 7-point Likert scale. Treatments are rated in terms of the degree to which they (a) are logical, (b) are scientific, (c) are complete, meaning helpful to all types of people who are depressed, (d) are likely to be helpful to an individual in other areas of life, (e) are likely to be considered as an option by the respondent if they were depressed, (f) are likely to be effective for most people, (g) require work from the client, and (h) are stressful for the client. In their first study using this measure, 252 undergraduate students rated written
descriptions of the treatment rationales for nine different therapies for depression. Item responses were factor analyzed, with the first seven items above forming a credibility factor, most often used as a total score in subsequent studies of credibility. The scale is useful across a variety of treatment modalities and treatment targets, and has been used extensively in a series of studies investigating treatments for PTSD (Feeny, Zoellner & Kahana, 2009; Zoellner, Feeny, Cochran & Pruitt, 2003; Zoellner, Feeny & Bittinger, 2009).

The credibility scale developed by Rokke et al. (1990) has been used in several studies to compare cognitive and behavioral treatment rationales as treatments of depression, with the latter treatment most often represented as activity change therapy, an early predecessor of modern behavioral approaches (Lewinsohn, 1974). Results across several studies have not been wholly consistent: in the development study for this measure (Rokke et al., 1990), participants rated activity change therapy as less credible than cognitive therapy, whereas this difference was only evident in younger adults in a study comparing younger and older adults’ credibility ratings (Rokke & Scogin, 1995). Likewise, in a clinical sample of depressed or dysphoric older adults, no differences in credibility were detected between a treatment focusing on thought changes or one focusing on behavior changes (Rokke, Tomhave & Jocic, 1999). A main research question of the present study addresses the need to clarify the inconsistent results obtained across several studies by Rokke and colleagues in terms of credibility ratings of cognitive and behavioral interventions.
Reason-giving and Differential Response to CBT Interventions

Prior to investigating agreement with the treatment rationale, Addis and Jacobson (1996) investigated the relationship between a client’s beliefs about the cause or causes of their depression and response to treatment in the CBT dismantling trial. They theorized that a potential mismatch between an individual’s belief about the cause of depression and the underlying model or theoretical mechanism of a given treatment might influence treatment outcome (potentially by influencing client engagement). Thus, they focused on individuals treated in the BA-only condition (n = 50) and the full CBT condition (n = 48), due to the distinctiveness of the treatment activities and rationales.

Patients in the dismantling study had completed the Reasons for Depression Inventory (RFD; Addis, Truax, & Jacobson, 1995), a 48-item self-report measure consisting of a collection of potential causes of depression (e.g., “I am depressed because…I think about things in a depressing way”) prior to beginning treatment. Responses are made on a 4-point scale, ranging from “definitely not a reason” (-2) to “definitely a reason” (+2), with subscales reflecting different broad categories of responses (e.g., Characterological, Biological, Childhood). The authors developed specific hypotheses with respect to these categories, but these were largely not supported and will not be discussed in detail here. Addis and Jacobson also specifically hypothesized that a pattern of responding consistent with what they called “reason-giving” – characterized by strongly endorsing multiple conceptualizations of depression, as captured by total RFD score – that might be reflective of a tendency toward a complex view of the disorder and its causes. This perspective is consistent with the widely-held premise that depression is caused by thoughts and feelings, which may tacitly reinforce
the notion that depression can only be cured through the similarly complex task of changing thoughts and feelings (see Hayes & Wilson, 1993). Such a perspective might suggest that certain ways of treating depression might be ineffectual because they are insufficiently complex; this might include behavioral interventions that emphasize the straightforward rationale of increasing contact with reinforcing experiences. In support of this notion, prior research has shown that ruminators induced to ruminate on the subject of causes of depression were less likely to report willingness to engage in pleasant activities, even though they rated them as being a potentially helpful way to improve low mood (Lyubomirsky & Nolen-Hoeksema, 1993). Thus, Addis and Jacobson theorized that individuals who responded in a “reason-giving” fashion on the RFD (i.e., higher total scores) should be expected to have poorer outcomes in the BA treatment conditions, due to potential concerns about this cause-treatment mismatch.

In this study, response to treatment was assessed by a composite of post-treatment BDI and HRSD while controlling for a comparable intake variable. Additional outcomes of interest included perceived helpfulness of the treatment (per self- and therapist-report at the 3rd session), along with more general client perceptions of the degree to which their assigned treatment matched with their ideas about what helps people in psychotherapy. Initial comparisons from the 69 patients included in the study suggested no differences in clients’ ratings of the degree to which either BA or CBT helps in the treatment of depression. Client perceptions of treatment helpfulness were significantly predictive of outcome in the BA condition \( r = .47, p < .01 \) but not in CBT. With respect to reason-giving, there was a robust negative association between RFD total scores (the proxy for reason-giving) and outcome in BA treatment \( r = -.38, p < .01 \) and no comparable
relationship in CBT. Thus, there was evidence that individuals who endorsed more reasons for depression – which the authors suggested is reflective of certain attitudes about how depression is best treated – were less likely to show positive outcomes in the BA condition, with no such relationship observed in the CBT condition.

**Reason-giving and Reactions to the Treatment Rationales**

Taken together, the two follow-up studies to the CBT dismantling trial (Jacobson et al., 1996) conducted by Addis and Jacobson (1996; 2000) suggest that patient beliefs about the cause and appropriate way of treating their depression influence their likelihood of responding to treatment. Treatment assignment moderated the relation between RFD and outcome in the behavioral condition, but not the full CBT condition. This finding in particular suggests the possibility that patients may respond differently to cognitive and behavioral interventions (or at least on the basis of their rationales). According to Addis and Jacobson’s explanation, this could be due to perception of a mismatch between the complexity of their problem, and the relative simplicity of the solution posed by behavioral interventions, or the degree to which they agree with the rationale of the treatment they are receiving.

Extending research on reason-giving outside of a clinical context, a follow-up study conducted by Addis and Carpenter (1999) evaluated the relationship between this construct and perceptions of treatment rationales. Fifty-one community volunteers were asked to read written descriptions of treatment rationales for activation-oriented (AO) and insight-oriented (IO) treatments for depression. Relevant to the present study, the AO rationale was consistent with the rationale for behavioral interventions used in CBT, and similar to the rationale described by Addis and Jacobson (2000) from the BA condition of
the dismantling trial. Participants provided ratings of these rationales on the aforementioned credibility scale (Rokke et al., 1990). After reading these written rationales, participants were shown videotaped vignettes of a mock client depicting her current symptoms and concerns, and then separate clips showing typical activities in each of these treatments (e.g., focusing on increasing pleasurable activities in the activation-oriented treatment). These vignettes provided a more naturalistic depiction of the rationale, along with interactions and activities typical of each treatment. The vignettes were rated on a measure assessing personal reactions to the rationales (PRR), using a 7-point scale in response to the degree to which the therapy presented: (a) would be helpful for the participant if he or she was depressed, (b) would help the participant understand the causes of his or her depression, (c) would help the participant learn how to cope with feeling depressed, (d) would be likely to be chosen if the participant was depressed, (e) would be effective in treating the participant’s depression. Summed scores of the credibility and PRR scales were moderately to highly correlated with each other within the same treatment (for AO, \( r = .52 \), for IO, \( r = .62 \)). Participants also completed the RFD and BDI, which were significantly correlated\( (r = .76) \).\(^{i}\)

RFD total scores were significantly predictive of both credibility and PRR ratings of the AO rationale and vignette (covarying initial BDI, partial \( r \) scores were -.31 and -.35 for credibility and PRR ratings respectively). No such relationships were evident with respect to the IO rationale. As expected, higher RFD scores were predictive of lower PRR and credibility ratings for the AO rationale and vignette. These findings are in line with those of Addis and Jacobson (1996), with reason-giving predicting less
positive evaluations of behavioral or activation-oriented treatments and worse clinical outcomes in treating depression.

The relationships between reason-giving, ratings of treatments, and treatment outcome, have not been investigated in the context of cognitive and behavioral interventions delivered separately. Prior research has compared these two approaches in terms of credibility and personal reactions, but not in conjunction with RFD ratings. Similarly, though reason-giving was associated with superior outcomes in the BA condition of the dismantling trial, participants in the CBT condition received both cognitive and behavioral interventions as opposed to an entirely cognitive approach. Further study is required in order to better understand how reason-giving and credibility are related to one another and to outcomes in the context of cognitive and behavioral interventions.

**Attitudes toward Treatments and Preferences**

Reflecting on the studies highlighted in the preceding sections, there appear to be meaningful associations between both reason-giving tendency and agreement with the treatment rationale, and outcomes in the treatment of depression. Reason-giving in particular was differentially associated with response to BA treatment, and ratings of treatment credibility and personal reactions to BA-type interventions. Credibility has been suggested as analogous to a more comprehensive, multi-item assessment of agreement with the treatment rationale, offering a potential connection between the two studies conducted by Addis and Jacobson (1996; 2000) and also linking these ideas to research on attitudes toward treatment options involving non-clinical populations (e.g., Rokke et al., 1990; Addis & Carpenter, 1999). The personal reactions measure developed
by Addis and Carpenter (1999) is also potentially informative, as it frames evaluation of treatments in terms of an individual’s own experiences and expectations, as opposed to the more general perception of credibility assessed by the Rokke et al. (1990) measure.

These findings suggest another potentially important avenue for investigation in the comparison between cognitive and behavioral interventions. If patients in the dismantling study had been told prior to starting the trial that they could choose between the treatment options, one might expect them to express preferences on the basis of their evaluation of the different conditions. For instance, patients high in reason-giving might be expected to reject the BA condition on the grounds of its incompatibility with their view of depression. Patient preferences have been the subject of extensive study in mental health research for years, and were recently identified as being a “demonstrably effective” way of adapting psychotherapy to improve outcomes, on the basis of expert panel reviews of the number of supportive studies, consistency and magnitude of positive results, experimental rigor of the studies, and other key dimensions (Norcross & Wampold, 2011). However, recent meta-analyses have emphasized methodological issues in this area of study, including in the assessment of preference itself.

Preference-outcome relations have not previously been studied in conjunction with other assessments of treatment options (i.e., credibility). In contrast to ratings of preference, credibility and PRR ratings consider multiple dimensions of a given treatment, including but not limited to an individual’s willingness to receive that treatment. Thus, while these constructs might be related, they may also differ with respect to their ability to predict differential treatment outcomes. The present study investigates the relationship between these constructs – specifically, the degree to which
preference can be predicted from credibility and PRR ratings – and then compare them as predictors of treatment outcomes known to be predicted by preference. The sections below provide a brief summary of the existing research on preference-outcome associations in depression, emphasizing in particular studies germane to the central idea posed by the proposed study: the comparison between cognitive and behavioral interventions within CBT for depression.

**Patient Preferences and Clinical Outcomes in the Treatment of Depression**

Preference has generally been conceptualized as constituting the behaviors or attributes of the therapist, or the therapy itself, that clients value or desire (Arnkoff, Glass & Shapiro, 2002). Preference is conceptually distinct from expectations, which relate more to beliefs about what should or is likely to happen in treatment (Swift et al., 2011), but in practice, the relationship between these concepts is more difficult to distinguish. Preference is further distinguished conceptually from treatment choice, though many studies of preference either explicitly or implicitly reflect choice. This review focuses primarily on treatment preferences, referring to patients’ inclinations toward one or another specific intervention type (e.g., preferring an insight-oriented treatment versus a behaviorally-oriented treatment).

In clinical research, treatment preference is most often measured by simply asking patients directly which intervention they prefer; in most studies, this question is asked prior to exposing patients to treatments or providing extensive information about them. More comprehensive measures and interviews have recently been developed in an effort to index different types of preference, but these are not yet widely adopted or well validated (Swift et al., 2011). In general, the virtue of these measures is that they ask
multiple questions instead of single items, and also typically evaluate the strength of a preference. This practice is regarded as important to understanding preference/outcome relationships, but is also not widely utilized in the field (Swift et al., 2011). The distinction between simple preference judgments and more elaborate evaluations of preference is one factor that appears to distinguish the vast majority of clinical research populations from studies involving laypersons or treatment analogues. In these latter samples, patients are provided with some detail about the nature of the treatment options – for instance, treatment rationales and sometimes depictions of treatment (e.g., Addis & Carpenter, 1999) – and then asked to make judgments. Paradoxically, patients who are actually eligible to receive these treatments may get less information about them, and provide less detailed responses about their reactions to these treatments. However, there have been no empirical tests of the degree to which this approach to assessing preferences diminishes the ability to detect associations with treatment outcomes.

In a meta-analysis of 26 studies and more than 2300 patients, Swift and Callahan (2009) reported that individuals receiving their preferred treatments were approximately half as likely to dropout prematurely, versus clients who were not matched to preferences or did not have their preferences considered. Clients receiving preferred treatments were also approximately 58% more likely to show improved outcomes, with a small $r$-type effect size of .15 (95% CI: .09 - .21). The collection of studies Swift and Callahan reviewed was diverse, with little overlap on outcome measures, even within a given diagnostic group. Effect size estimates were significantly heterogeneous (per meta-analytic analyses) and varied widely from -.26 to .55, with negative signs indicating an advantage for the non-preferred treatment. The authors noted that the way preference
was related to treatment assignment had an important influence on reported outcomes.
Preferences are the focal point of match/no-match studies and determine which treatment patients will be assigned to, whereas in standard randomized clinical trial (RCT) designs, preferences are assessed but do not influence randomization. Partially randomized preference trials (PRPTs) attempt to blend these approaches by allowing patients with strong preferences to choose their own treatments, and randomizing individuals without strong preferences. As might be expected, preference effects are notably weaker in these designs (Swift & Callahan, 2009; Swift et al., 2011), presumably because of the reduction of individuals with strong preferences who are assigned to a non-preferred treatment. This finding emphasizes the importance of assessing strength of preference, as studies that fail to do so may underestimate important preference-outcome relations that are not evident in a simple categorical preference rating.

A follow-up study on preferences conducted by Swift et al. (2011) was expanded to include 35 studies, and reported similar findings: namely, preference-matched clients were approximately half as likely to drop out, and showed greater improvements in outcomes (with a $d$ of .31, corresponding to a small effect). Studies of preference-outcome effects in depression showed an average effect size of $d = .35$ in terms of the magnitude of difference between patients matched to preferred versus non-preferred treatments. This most recent meta-analysis contained nine studies that report on preference-outcome associations in the treatment of depression, with an additional study (Kwan, Dimidjian & Rizvi, 2010) that was not included in the meta-analysis bringing the total number of relevant studies to ten. To briefly summarize this literature, Table 1 provides a list and a brief description of the key features of these studies. Notably,
outcome measures varied considerably, both in terms of actual measures used (e.g., HRSD, BDI) and general domain of interest (i.e., therapy process, symptom change, response/remission or dropout). The mean study size was 172 patients, with a range of 22 to 429. Individual preference-outcome Cohen’s $d$ effect size estimates ranged from -.18 to 1.15.

Table 2 summarizes the findings with respect to preference-outcome relationships. Nine studies explored symptom change as a predicted outcome, with only one study finding a clear preference-outcome relationship (Kocsis et al., 2009), and four doing so in a partial fashion (i.e., for one treatment type; at mid-point but not post-treatment; for a secondary measure; mediated by attendance). Four studies had relapse, response or diagnostic criteria as dependent variables, but only Kocsis et al. (2009) found evidence of a preference-outcome association. Therapy process variables were predicted in four studies, with one finding clear evidence of an association between preference and therapeutic alliance, one finding no effect, and the remaining two finding partial evidence of preference effects (i.e., willingness to augment treatment with medication; preference strength but not match predicting adherence to treatment). Finally, attendance and dropout were examined in six studies, with three finding preference-dropout associations, and three failing to find any association. In conclusion, dropout and symptom change measures appeared to be the most reliable outcomes associated with preference in these studies. These outcomes were also among those most commonly examined in the studies comparing treatments relevant to the proposed project, and will therefore be treatment outcome targets in this study.
Two studies included in the above collection bear special consideration as they specifically consider cognitive and behavioral approaches to treating depression. Lamentably, both studies fail to include information that might be influential in determining how best to evaluate preferences for these two intervention types. Kwan et al. (2010) explored treatment preference-outcome effects in an RCT involving 106 depressed outpatients treated with medication, behavioral activation therapy (BA) or CBT. Patients were asked a simple preference question with options of psychotherapy, medication, or no preference; unfortunately, no distinction was made for preferences for BA or CBT. Patients who received their preferred interventions were less likely to refuse the results of their randomization, showed bettered attendance and were less likely to drop-out of treatment, but preferences did not predict the therapeutic alliance early in treatment. Change in symptoms (as assessed by the slope of HRSD and BDI scores across treatment) was not directly related to preference, but there was evidence of a relationship mediated by attendance. Critical limitations of this study include a strong preference bias for psychotherapy, in conjunction with an atypically high dropout rate in the medication condition (50%).

Rokke et al. (1999) explored patient preference in the form of choice in a match/no-match study of 10 sessions of self-management therapy for 40 depressed older adults. Patients were randomly assigned to a condition where the treatment focus was chosen by the patient (and therefore assumed to be preferred), one where the treatment focus was assigned, or to a waitlist. Patients who were allowed to choose could focus on either behavior or cognitions as targets for change; the rationales described for these treatments closely mirror those for cognitive and behavioral approaches used in CBT, but
with a more didactic approach to skill-learning. Rokke and colleagues also asked participants to rate these treatment options using the credibility scale described earlier (Rokke et al., 1990).

This study differs from most others included in the Swift et al. (2011) collection for its focus on the importance of choice rather than preference. Regrettably, as a result, Rokke et al. (1999) provided no information about the relationship between credibility ratings and choices made by subjects, nor did they consider the preferences of individuals randomized to the assigned-target condition. Furthermore, relationships between credibility scores, patient preferences, and treatment outcomes were not described. However, they did note that there was no evidence of an overwhelming preference for treatment targets (58% of patients chose a cognitive focus) and no difference in credibility ratings of the two approaches. Individuals who chose their therapy target were not more likely to show symptom improvements but were less likely to drop out (after removing individuals who left the study due to extenuating circumstances) and showed a trend toward completing more sessions.

It merits comment that a series of studies in the PTSD treatment literature have provided a comprehensive exploration of the relationships between treatment credibility, PRR ratings, and preference for common treatments for PTSD (typically, prolonged exposure, sertraline or their combination). The collection of studies by Zoellner, Feeny and colleagues (Cochran, Pruitt, Fukuda, Zoellner & Feeny, 2008; Feeny et al., 2009; Pruitt, Zoellner, Feeny, Caldwell & Hanson, 2012; Zoellner et al., 2003; Zoellner et al., 2009) has examined the relationships between these constructs in both clinical populations (i.e., treatment-seeking PTSD patients) and analogue samples (typically,
undergraduate women with a history of exposure to trauma), and has investigated both written and videotaped depictions of treatments and treatment rationales. These studies have generally found that credibility and PRR ratings are associated with treatment preferences.

**Purpose of the Present Study**

The preceding review has established several ideas that serve as the foundation to the present study. First, there is a need to better understand how CBT for depression achieves its effects, and a comparison of cognitive and behavioral interventions is one potential way to do so. There is further reason to believe that individuals vary in their beliefs about the reasons for their own depression, as well as the degree to which they regard cognitive and behavioral interventions as credible and personally viable treatments. It may follow that these beliefs extend to influence their preferences for treatment, which show a small but consistent association with treatment outcomes (notably, symptom change and dropout).

Relationships have been identified between some of these constructs (i.e., reason-giving and credibility), and there is evidence that these measures may predict certain treatment outcomes (i.e., reason-giving and symptom change; preference and symptom change / dropout). However, additional research on the nature of these relationships is warranted in the context of a comparison between cognitive and behavioral interventions when delivered separately. Specifically, this project seeks to understand the relationship between reason-giving and credibility ratings of these two interventions - a relationship that has not previously been examined. An additional aim is to investigate whether reason-giving tendency and credibility ratings predict treatment preference, an individual
difference variable associated with treatment outcomes. In turn, this may represent an avenue for prediction of differential response to treatment in the context of the study comparing cognitive and behavioral interventions, and to extend these findings to empirical tests of these relationships in CBT in general.

Four main research questions guided this project. The first question concerned whether cognitive and behavioral interventions are regarded as equally credible treatment options. To test this, written treatment descriptions and video vignettes depicting cognitive and behavioral interventions were developed in order to provide a context for depicting rationales and common activities of such treatments (following the method of Addis & Carpenter, 1999). Prior to data collection, a group of graduate therapist trainees familiar with CBT rated the two vignettes as comparable on a number of dimensions (described below).

The written descriptions and vignettes were rated using the credibility (Rokke et al., 1990) and PRR scale (Addis & Carpenter, 1999) described above. Two groups provided ratings: a large, analogue sample of dysphoric undergraduates, and a smaller sample of depressed outpatients recruited from the clinical trial of behavioral and cognitive interventions. These two groups were selected for several reasons. Data from the outpatient sample allowed for an investigation into how the beliefs and opinions that individuals hold may predict differential treatment response. As there is also a need to understand what motivates an individual’s willingness to enter treatment, credibility and PRR ratings were examined in a group of dysphoric undergraduates who were not uniformly seeking treatment. The inclusion of this group provided a broader range of opinions, as the outpatient group had already agreed to participate in a trial of
psychotherapy and is therefore unlikely to include individuals with strongly negative attitudes toward either or both of these treatments. Practically, recruitment of dysphoric undergraduates is easier on a larger scale, which provided greater statistical power for analyses. Dysphoric samples are often used in depression research when clinical populations are not readily available (e.g., Lowenstein & Hokanson, 1986).

Following the example of Addis and Carpenter (1999), ratings of personal reactions to the treatment rationales (PRR) were included in order to evaluate potential differences in individuals’ beliefs about their own treatment options versus their general sense of treatment credibility. However, as the PRR has been used less frequently, the main study hypotheses focused on credibility ratings, with PRR analyses described in a separate section of this paper.

Research Hypotheses

Prior research in which cognitive and behavioral treatments have been compared using the credibility scale has yielded mixed results (Rokke et al., 1990; Rokke & Scogin, 1995; Rokke et al., 1999). Therefore, Hypothesis 1 was that cognitive and behavioral interventions would be comparably credible in terms of the overall ratings provided by participants. From an analytic standpoint, credibility ratings (total scores) were compared as a function of intervention type. Post-hoc analyses on particular items from the credibility scale were to be conducted in the event that robust overall differences in credibility were detected.

The second research question concerned the relationship between reason-giving (as indexed by total score on the RFD) and credibility ratings of cognitive and behavioral interventions. As this question is connected to prior studies in both a clinical (Addis &
Jacobson, 1996) and non-treatment context (Addis & Carpenter, 1999), the relationship between these constructs was explored in both the dysphoric undergraduate and patient samples. Prior research strongly suggests an association between reason-giving and ratings of behavioral interventions, though there is no prior research involving ratings of purely cognitive interventions. Two hypotheses and corresponding sets of analyses were conducted. Hypothesis 2A was that RFD scores would predict credibility ratings of the behavioral treatment, but not ratings of the cognitive treatment; based on prior research, higher RFD scores were expected to be associated with lower behavioral credibility ratings. Implicit in the results observed by Addis and Jacobson (1996) is the notion that the full-CBT rationale is more credible to individuals high in reason-giving. Because the cognitive rationale presented in these vignettes is similar to the full-CT rationale presented in the dismantling study, it should be regarded as more credible than the behavioral rationale to reason-givers. Thus, Hypothesis 2B was that RFD total score would predict the relative difference in credibility ratings between the cognitive and behavioral rationales, with higher RFD scores associated with higher ratings of cognitive versus behavioral vignettes.

The third research question related to how RFD and credibility ratings were associated with measures of treatment preference. As noted above, these relationships have not previously been explored in the context of cognitive and behavioral interventions. While the research question was the same in both samples, slightly different tasks were utilized, due to inherent differences between the two groups. In the dysphoric group, participants were asked to consider which treatment they would seek if they became depressed; they provided responses indicative of their preferences on a 7-
point valence scale, with responses ranging from no preference to strong preference for
cognitive or behavioral in order to capture both valence and strength of these judgments.
A slightly modified version of the task in the sample of clinical patients reflected an
assessment of their preferences for the treatments they were eligible to receive. RFD
total score and the credibility difference score described above were evaluated as
predictors of preference ratings, both separately and together in the same regression
analyses. Hypothesis 3 was that both RFD and credibility difference scores would be
predictive of preference ratings, with both scores remaining significant predictors in a
comparative model. As described above, both RFD total scores and credibility difference
scores should be associated with a relatively strong preference for cognitive
interventions.

The fourth research question related to whether reason-giving, credibility ratings,
and preference predict clinical outcomes in depressed outpatients, including differential
response to cognitive and behavioral interventions. These analyses were limited to the
outpatient sample. Based on the results of the Swift et al. (2011) meta-analysis, and
further examination of trials involving depression and CBT treatments, two main
outcomes were predicted: symptom change and dropout rate. These outcomes were
examined across the first eight weeks of treatment. RFD total score, credibility
difference score, and preference ratings were evaluated as predictors of dropout and
symptom change. Dropout models were conducted using logistic regression, while
symptom change models were conducted using hierarchical linear modeling (HLM) using
the HRSD as a primary symptom change variable. As these analyses were primarily
concerned with differential response, interactive effects were tested using the products of
the predictor variables and a dummy coding variable reflecting treatment assignment (behavioral/cognitive). Predictor variables were evaluated independently and in multi-predictor models where indicated. Basic descriptive statistics related to treatment outcomes and treatment assignment were also generated prior to conducting analyses.

There were several hypotheses related to this research question. Hypothesis 4A was that the interaction of treatment assignment and RFD total score would predict dropout and symptom change, with this interaction driven by the association between RFD total score and outcomes in the behavioral condition. Hypothesis 4B was that a significant interaction between credibility difference scores and treatment assignment would be identified, such that symptom change would be greater and dropout risk would be lower when treatments were rated as more credible. Finally, preference by treatment assignment interactions were also evaluated as predictors of dropout and symptom change. Hypothesis 4C was that individuals who receive their preferred intervention would show reduced risk of dropout and superior symptom change; for these analyses, both continuous ratings of preference and a categorical treatment match variable were considered. Exploratory analyses comparing the relative predictive ability of these variables (i.e., RFD total, credibility difference score, and preference score) in a multi-predictor model are also reported.
Chapter 2: Methods

Data Collection and Cleaning

All research activities were reviewed and approved by the institutional review board of the Ohio State University. Two groups were recruited for this study: a sample comprised of dysphoric undergraduate students (hereafter, the student sample), and a smaller sample of patients participating in a pilot study comparing cognitive and behavioral interventions (the patient sample). Descriptive statistics are provided below, after a brief explanation for the removal of certain participants and the resulting sample sizes.

Before beginning primary analyses, patient and student data were examined for completeness and accuracy. Participants in the patient sample completed most measures on paper, with the exception of ratings of the interventions, which were conducted via an online survey website. Participants in the student sample completed all measures electronically.

Both groups had members with partial and incomplete data. When data issues affected ratings of the interventions, participants were considered ineligible for inclusion, and were therefore removed for one of three reasons. First, some participants did not complete ratings (typically due to failure to complete the experiment, and less often due to skipping of multiple vignette questions). Second, participants who incorrectly identified one or both vignettes were removed (e.g., identifying the cognitive vignette as
behavioral, or vice versa). Note that participants were provided with information that correctly identified each vignette at multiple points during the experiment, including at the end of each video. Finally, participants could be removed for demonstrating a failure to comprehend vignette content, as evidenced by incorrect responses to a substantive majority (i.e., at least 4 of 5) of simple comprehension questions, that were asked following presentation of each vignette. In the student sample, 196 of 269 participants (73%) had data eligible for inclusion, while data were available for 38 of 42 patients (90%). As noted below, for some descriptive statistics, sample sizes vary due to missing responses on individual items.

**Student Sample**

The student sample was recruited through the Research Experience Program (REP) of the Psychology Department at the Ohio State University, and received course credit for participating. Students in this sample scored 9 or above on the BDI (Beck et al., 1961) during a pre-screening assessment conducted through REP, which is conventionally regarded as evidence of dysphoric mood (e.g., Alloy, Fedderly, Kennedy-Moore & Cohan, 1998). Participants with pre-screening BDI scores in this range were allowed to access the experiment through the REP site, and could choose to participate. After enrolling, they were contacted by email with a link to complete the study.

**Demographics & descriptive statistics.** All statistics reported herein are based on the full student sample minus one participant who did not complete demographics questions (N=195, except where noted). The mean age of the student participants was 19.1 (SD = 2.5, range 18 – 41), and approximately three-quarters were female (n =143, 73.3%). In terms of ethnicity, this sample was approximately 76% Caucasian (n=148),
11% Asian (n=22), 4% African-American (n=7), 1.5% Hispanic (n=3) and 8.8% participants who identified as multiracial (n=15). The undergraduate sample was overwhelmingly unmarried (n=192, 98%) and 68% were in their first year of college (n=132).

Symptom data for the PHQ-9 and BDI-II were available for all eligible participants in the student sample, with BDI-II scores re-scaled for comparison purposes (see note below). The mean BDI-II score was 15 (SD=10.6; range 0 -60), and the mean PHQ-9 score was 7.9 (SD = 5.7, range 0 – 25), reflecting mild depression. These scores were highly correlated (r = .85, p <.0001). Using severity criteria developed by the authors of the PHQ-9, 36% of student participants reported scores consistent with no depression (n=70), 31% reported mild depression (n=60), 21% reported moderate depression (n=41), 7.7% reported moderately severe depression (n=15) and 5% reported severe depressive symptoms (n=10). The PHQ-9 can also be used to assess provisional MDE status, based on a comparison to DSM-IV-TR diagnostic criteria (as described in Kroenke, Spitzer & Williams, 2001). Using this scoring formula, approximately 16% (n=32) of the sample would be likely to meet current criteria for a major depressive episode.

Considering experience with and current symptoms of mental health issues, approximately 20% of the student sample reported that they had experienced a mental health issue at some point in their life (n=40), with 9% reporting prior personal experience with depression (n=18). Eighteen percent of the student sample reported prior therapy or counseling (n=35), and 11% (n=22) said they had previously received medication for a mental health issue.
**Patient Sample**

The outpatient sample was comprised of patients enrolled in a pilot study comparing cognitive and behavioral interventions in the treatment of depression. Participants were recruited through advertisements and word-of-mouth from the central Ohio region. They were pre-screened by phone prior to being assessed to ensure likely fit with study characteristics and a general interest in participating, then invited to schedule an intake assessment as appropriate.

**Study assessments & inclusion criteria.** Intake procedures included the Structured Clinical Interview for the DSM-IV for both Axis I disorders (SCID; First, Spitzer, Miriam, & Williams, 2002) and Axis II disorders (SCID-II; First, Spitzer, Gibbon, Williams, & Benjamin, 1994), and the Hamilton Rating Scale for Depression (HRSD; Hamilton, 1960). Inclusion criteria for the treatment study were (a) current diagnosis of Major Depressive Disorder according to DSM-IV criteria (APA, 1994) based on the SCID interview, or current Major Depressive Episode in the context of a Bipolar II diagnosis; (b) being 18 years of age or older; and (c) willingness and ability to provide consent. Exclusion criteria included (a) a history of Bipolar I disorder or psychosis; (b) current substance dependence within the last 6 months; (c) a current Axis I disorder that is clearly the predominant clinical issue and which requires treatment other than that being offered; (d) current suicide risk or significant self-harm sufficient to preclude outpatient treatment; (e) subnormal IQ; or (f) clear indication of secondary gain (e.g., court-mandated treatment, monetary incentives). 83 patients were invited to an intake assessment; 23 of these declined to schedule or did not show up for a scheduled intake appointment. 18 participants did not qualify, for reasons that included: substance
dependence (n = 6), psychotic disorder or Bipolar I history (n = 5), failure to qualify for current major depressive episode (n = 3), or recent medication change (n = 4).

After qualifying and consenting to participate in the trial, participants were randomly assigned to eight weeks of treatment with either a cognitive or behavioral focus. They were assessed with the HRSD approximately four weeks after their last intake appointment (at the midpoint of their first module), and then again at the 8-week midpoint of treatment. Participants were then re-randomized to either cognitive or behavioral focus (resulting in continuation or switching of treatment types). Data collected after the midpoint of treatment were not considered in the present study. Qualifying participants received free psychotherapy and were eligible to receive a small financial incentive (i.e., gift card) for completing the final study assessments.

**Eight-week outcome interval.** Analyses involving the pilot treatment study focus on the first eight weeks of treatment (i.e., the first treatment assignment). This eight-week window includes symptom assessments at the start, midpoint and end of this time period, allowing for multiple ways of evaluating symptom change. This window also allows for a more straightforward test of intervention type (without the added complication of a potential switch in treatment focus in the second half of treatment). Furthermore, there are several reasons to think that this time period will allow for a reasonable test of these associations. First, early, rapid change is common in standard CBT (Ilardi & Craighead, 1994), and differential response to CBT has been identified in as little as two weeks (Fennell & Teasdale, 1987). Patients whose symptoms are sufficiently severe may have completed as many as 12 sessions in this time window, which is ample time to have begun working on typical treatment interventions. Second, prior studies of preference-
outcome associations have detected effects in shorter time windows. For instance, Elkin and her colleagues (Elkin, Yamaguchi, Arnkoff, Glass, Sotsky, & Krupnick, 1999) showed that preference match affected dropout rate in the first four weeks of the Treatment of Depression Collaborative Research Program. Finally, two prior trials conducted in the Depression Treatment and Research Clinic at Ohio State have had elevated dropout rates in the first eight weeks relative to the latter half of treatment.

**Demographics & descriptive statistics.** Forty-two participants met inclusion criteria for the study; however, as noted above, only 38 participants had all data necessary for inclusion in these analyses. The sample was comprised of 17 males and 21 females, with a mean age of 33.2 years (SD = 15.2, range 18-65). Four participants were older than 60 at the time of their intake. The sample was predominantly white (n=31; 82%), with three African-Americans and four participants describing themselves as being multi-ethnic. In terms of marital status, 79% of the sample was currently single (n=30). Eleven participants reported full-time employment, while ten participants were working part-time (less than 40 hours a week). Fourteen participants (50% of the sample) were students. Two participants indicated that they were unemployed but not expected to work due to disability, and the remaining seven participants (18%) were unemployed and looking for work. A third of the sample (n=12; 32%) reported having completed college or advanced degrees.

**Diagnoses and symptoms at intake.** All participants met criteria for a current major depressive episode at the time of their assessment; one participant met criteria for Bipolar II disorder, with a current major depressive episode; all others met for Major Depressive Disorder, per the SCID-I assessment. BDI-II scores at intake were collected
from 35 subjects, with a mean score of 35.1 (SD = 9.1 range 18-52). The mean HRSD score at intake was 20.2 (SD = 5.2, range from 8 – 31) indicating moderate depression. Nine patients (24%) had “double depression”, characterized by concurrent dysthymia and major depressive episode diagnoses. Twenty-two patients (58%) endorsed recurrent depression (between one to five or more past episodes), while sixteen patients (42%) were identified as currently experiencing chronic depression.

In terms of comorbid diagnoses, 10 patients reported no comorbid conditions, while the remaining patients ranged from one to four additional Axis I diagnoses. Anxiety comorbidity was most common, with 23 patients having at least one comorbid anxiety diagnosis, most commonly GAD (n=17). Twenty-one patients had at least one personality disorder, most often from Cluster C (n=13), most commonly avoidant PD or obsessive-compulsive PD. Seven patients had multiple Axis II diagnoses, ranging in number from two to five.

**Psychotherapy Study Treatment & Conditions**

**Therapists & training.** Six trainee therapists (three women) conducted treatment and assessments. Their past experience with CBT varied, but all therapists received approximately 40 hours of specific training on the separate implementation of cognitive and behavioral interventions, including the use of intervention-specific techniques and common elements of CBT. Therapists completed measures of intervention fidelity after every session, and there was no self-reported indication of protocol breaches (e.g., therapists conducting behavioral interventions with a client assigned to cognitive treatment).
**Treatment and therapist assignment.** Patients were randomly assigned to behavioral or cognitive interventions after completing their intake assessment. They were also randomly assigned to a trainee therapist using a weighting paradigm that considered each therapist’s availability for clients and their relative proportion of cognitive and behavioral clients, to ensure that all therapists treated comparable numbers of clients with both interventions. There were equal numbers of patients assigned to behavioral and cognitive in both the full sample (n=21 per intervention) and the sample with vignette rating data (n =19 per group). Therapists saw between four and eight clients each (one therapist was only carrying a half case-load), with comparable numbers of clients in both behavioral and cognitive interventions (e.g., 3 cognitive / 4 behavioral).

**Format of cognitive and behavioral interventions.** The cognitive intervention arm focused on the basic cognitive model of depression, predominantly utilizing those interventions described in the AT arm of the Jacobson et al. (1996) trial (minus those included in the BA arm of that trial). Therapists were instructed to focus on identifying and scrutinizing automatic thoughts, considering evidence and alternative explanations, and using hypothesis testing to challenge negative thoughts. Therapists were not prohibited from discussing schema-related issues (unlike the AT arm of the dismantling trial) but were not obligated to do so. Behavioral techniques were proscribed (e.g., activity logging, assertiveness training) but therapists were allowed to work flexibly within a cognitive framework to address problems that might otherwise be addressed by behavior changes (e.g., gathering evidence for an alternative perspective by modifying behavior). Thought records comprised the majority of homework assignments in this condition.
The behavioral intervention condition emphasized a behavioral rationale, focusing on establishing reinforcing connections to encourage pleasure and mastery experiences, identifying and correcting problematic behavioral patterns, and augmenting social skills. Likewise, cognitive interventions were proscribed (e.g., thought records), and therapists were encouraged to flexibly employ behavioral principles to handle negative thoughts (e.g., considering the form rather than content of thoughts). In designing the behavioral condition for this study, every effort was made to model the intervention after the behavioral elements of CBT (akin to the comparable condition of the dismantling study) as opposed to adopting elements of more recently developed behavioral activation treatment (e.g., Martell, Addis & Jacobson, 2005).

Session frequency. Sessions were scheduled to occur weekly across the trial in both conditions, with patients showing higher initial severity (e.g., intake HRSD >20) recommended to start at twice weekly for the first several weeks, in accordance with the CBT manual (Beck et al., 1979). In practice, sessions did not always occur this frequently for all patients in the study. The mean number of sessions in the first eight weeks was 6.52 (SD = 2.5, range 1 – 12).

Cognitive and Behavioral Treatment Rationales & Vignettes

Vignettes depicting behavioral and cognitive interventions were developed based on the approach utilized by Addis and Carpenter (1999). There were three segments: an introductory segment based upon the description of the client in that earlier study (M. Addis, personal communication), and then one segment introducing each of the behavioral and cognitive approaches. The order of presentation of these latter two segments was randomized across participants. Participants were also presented with a
brief written rationale describing each of the interventions. These were based on previously published descriptions (Addis & Carpenter, 1999), and modified to reflect the focus of this study (intervention versus treatment). These were matched as closely as possible for length and number of words, and did not differ substantially on sentence structure, grade level, and reading ease as judged by Microsoft Word software (Microsoft Inc., 2007).

After clicking a link presented during the online experiment, videos loaded in a separate window for participants to view. All three video segments began with voice-over introduction of the basic topics, including the previous session’s homework in the intervention sections. The introductory segment was approximately three minutes long, and introduced the major issues the client with which the client had been coping (i.e., sleep, stress at work, self-criticism, missing her family, and reduced reinforcers). The intervention segments were each approximately eight and a half minutes, and depicted the client and therapist considering the client’s problems in a manner typical of the intervention being portrayed. In the behavioral segment, the client and therapist reviewed an activity log and noted the positive effects of incorporating physical activity to stave off low mood. They also considered managing stress at work by participating in a walking group at lunch, and addressed poor sleep habits that might contribute to tiredness. In the cognitive segment, the session was extensively devoted to a discussion of a thought record related to perceptions of work stressors, and included challenging thoughts by considering evidence, noting the probable effect on mood and coping that might come from re-appraising the situation, and considering rational responses in the event of similar future stressors.
**Vignette & rationale evaluation.** Every effort was made to make the two intervention videos as comparable as possible, with an ultimate goal of faithfully depicting typical and commonly-used exercises for both intervention types, rather than ensuring strict equivalence with respect to the problems being targeted. The videos were also designed to be fairly equivalent in terms of frequency of changing angles and the number of wide shots that include both therapist and client.

Graduate students (n = 6) with knowledge and experience in CBT were asked to rate the vignettes to ensure comparability on several dimensions including: basic information about the quality of the video presentation (e.g., sound, image, etc.); perception of benefit to client; quality of the alliance depicted in the video; client agreement and trust with therapist; how typical / representative this clip was of the intervention being portrayed; how typical the client’s issues were of depressed people; how severe / difficult / receptive the client was; how effective was the approach for these problems (as portrayed); and how much the client contributed in the clip. Ratings were made on a seven-point scale. Vignettes were generally regarded as easy to view and hear, and written comments indicated that they were comprehensible. Analyses comparing responses on specific items (via dependent t-tests) identified a single significant difference (\( p = .03 \)) on the question “How much trust was there between the client and the therapist?” However, while the difference between group means (0.67) was statistically significant, it was also small and comparable in magnitude to other items. These ratings therefore do not suggest any strong reason to suspect meaningful differences between the vignettes in terms of the dimensions described above.
Measures

Ratings of Cognitive and Behavioral Interventions

**Brief comprehension questions.** Participants were asked five, simple multiple-choice questions pertaining to the content of the vignettes as a safeguard against failure to follow study procedures. These questions were straightforward and reflective of the main subjects of conversations in each corresponding vignette. As noted above, participants were excluded if they showed poor comprehension (i.e., four or more incorrect answers).

**Intervention Credibility Scale (ICS).** As described earlier, the credibility scale (Rokke et al., 1990) relates to multiple dimensions of treatments thought to influence perceptions of credibility and to instill the potential for change. While the original scale had nine items, most studies have focused primarily on the first seven items, typically summed to form a credibility score. Answers are given on a seven-point Likert-type scale. Analyses reported by Rokke et al. (1999) from a prior study listed a Cronbach’s alpha of .89 for the first seven items on the scale. All participants completed this measure twice, one time after viewing each of the cognitive and behavioral treatment vignettes and reading the corresponding rationales. Cronbach’s alpha for the behavioral vignette was .84 for the student sample, and .83 for the patient sample. For the cognitive vignette, Cronbach’s alpha for the student sample was .85, and .84 for the patient sample.

**Personal Reactions to the Rationales (PRR).** The PRR is a five-item measure developed by Addis and Carpenter (1999) to evaluate the differences between general reactions to treatment rationales of different psychotherapies, and participants’ individualized reactions to the treatments relevant to their own treatment options. Answers are given on a seven-point Likert-type scale. All participants completed this
measure twice, one time after viewing each of the cognitive and behavioral treatment vignettes and reading the corresponding rationales. For the PRR, Cronbach’s alpha for the behavioral vignette was .89 for the student sample, and .85 for the patient sample. For the cognitive vignette, Cronbach’s alpha was .91 for both the student and patient sample.

**Treatment preference task.** Both samples completed therapy preference ratings at the end of the study, after completing all other questionnaires and watching all videos. A prompt asked participants: “What kind of psychotherapy (cognitive or behavioral) would you go into, and how strong is your preference?” As the dysphoric students were not formally diagnosed and many had low levels of depressive symptoms, they received an additional instruction preceding this prompt that read: “Try to imagine that you have been depressed for a while and decide to seek professional help.” Scores were scaled so that a preference for cognitive treatment resulted in a negative value, with -3 corresponding to “strongly preferred cognitive”, and +3 corresponding to “strongly preferred behavioral” (numbers were not actually presented on the screen). Participants could also select “no preference (either one)”, corresponding to a 0 on the scale. Additionally, to capture the possibility that participants might dislike both treatments, and struggle to select an option adequately depicting this, a separate question was asked of all participants: “If other options were available, I would not choose either cognitive or behavioral psychotherapy”.

**Symptom Measures & Diagnostic Assessments**

**Beck Depression Inventory (BDI).** The BDI (Beck et al., 1961) is a reliable and well-validated measure of depressive symptoms (Beck, Steer, & Garbin, 1988). It is a 21-item self-report scale with item scores ranging in value from 0 to 3, thus possible
scores are 0 (minimal depression) to 63 (high depression). This measure was used in REP prescreening to identify individuals with high and low depressive symptoms and was only completed by the student sample.

**Beck Depression Inventory - 2nd Edition (BDI-II).** The BDI-II (Beck et al., 1996) is the most current revision of the BDI (see above). It is a 21-item self-report instrument used to assess the existence and severity of symptoms of depression according to the Diagnostic and Statistical Manual of Mental Disorders - Fourth Edition (DSM-IV). Item range in value from 0 to 3, and possible scores are 0 (minimal depression) to 63 (high depression). The version implemented for the student sample in this study had removed the item pertaining to suicidal thoughts, so it contained only 20 items, reducing the maximum possible score to 60. For comparison purposes, BDI-II scores were rescaled to be comparable to those obtained from the treatment sample, by multiplying the total score by 63/60.

**Patient Health Questionnaire (PHQ-9).** The PHQ-9 (Kroenke, Spitzer, & Williams, 2001) is a 9-item self-report measure based on the diagnostic criteria for major depression from the Diagnostic and Statistical Manual Fourth Edition (DSM-IV). It has been used in primary care and research settings to assess depression severity, symptomatology, and functional impairment. A scoring guide allows for the assessment of depressive symptom severity, as well providing a formula to determine the prospective likelihood of meeting criteria for a current major depressive episode.

**Hamilton Rating Scale for Depression (HRSD).** The 17-item HRSD (Hamilton, 1960; Williams, 1988) was used as the primary measure of depressive symptom severity, modified to assess atypical symptoms (e.g., hypersomnia; Miller,
Bishop, Norman & Maddever, 1985). Clinical trainees administered this questionnaire to participants in the pilot sample at intake assessment, again approximately four weeks later, and again approximately eight weeks after intake. The HRSD has demonstrated adequate internal consistency (Cronbach’s $\alpha \geq .70$) and the ability to discriminate between levels of depressive symptoms (Bagby, Ryder, Schuller, & Marshall, 2004). A subset of ratings from across the entire trial (i.e., intake, week 8, post-treatment) were selected at random and re-rated to assess reliability. The intra-class correlation coefficient for these ratings was .98, reflecting very good reliability.

**Structured Clinical Interviews (SCID, SCID-II).** The SCID (First, Spitzer, Miriam, & Williams, 2002a, 2002b) and SCID-II (First, Gibbon, Spitzer, Williams, & Benjamin, 1997) for the Diagnostic and Statistical Manual of Mental Disorders (DSM-IV; American Psychiatric Association, 1994) were used to assess whether participants meet criteria for Axis I and II conditions, including current major depressive episode and other primary inclusion/exclusion criteria.

**Patient Characteristics Measures**

**Demographics questionnaire.** All participants were given a demographics measure to collect information about their age, ethnicity, sex and marital status for descriptive purposes.

**Reasons for Depression Inventory (RFD).** As described in detail above, the RFD (Addis, Truax & Jacobson, 1995) was developed to evaluate endorsement of different causes of depression. Ratings are provided on a 4-point Likert-type scale, with items reflecting various ways of conceptualizing the source of one’s depression, provided in response to the question “I am depressed because…” The original measure was
comprised of 48 items and was categorized into eight different subscales, with the total score on the RFD has also been used as a measure of “reason-giving”. A more recent study (Leykin et al., 2007) utilized a 14-item short form version of the RFD, with 13 items from the original scale and a new item added (“I was born to be depressed”) because of concerns about missing content. A prompting statement was added for the dysphoric sample to reflect the fact that many participants were not depressed, in keeping with modifications used by prior authors using non-depressed samples. This statement instructed participants who are not currently feeling depressed to consider a time when they were previously depressed or a time when they were extremely sad for more than a little while.

Ten patients and 17 students omitted one or more items on the RFD, so their total scores were recalculated to make them commensurate with other participants (using the method described for the BDI-II above). Comparison of mean values indicated significant differences between the patient sample (mean = 23.1, SD = 6.7, range 9 – 35) and the student sample (mean = 11.3, SD = 7.5, range 0 – 34) for total RFD scores ($t (232)= -9.0$, $p < .0001, d = 1.2$).

**Study Procedures.** Participants in the patient sample were first assessed (as described above) to determine eligibility for the study. Both groups completed questionnaires (e.g., RFD, demographics, etc.) prior to viewing the vignettes and rationales, and making preference ratings. Participants first viewed the 3-minute introductory vignette. Viewing order of the two intervention vignettes was randomized, with approximately half the participants seeing the cognitive intervention video and treatment rationale first. As an additional check on compliance and comprehension,
participants were asked to identify which video they had just watched. Credibility and personal reactions ratings were conducted after each of the two main vignettes. In the last step, participants completed the treatment preference task.
Chapter 3: Results

The results presented in this section are organized to mirror the hypotheses laid out in the introduction. Descriptive statistics are first presented, followed by results for the first three hypotheses in both samples. General findings pertaining to the pilot study are presented prior to results pertaining to the fourth set of hypotheses, which are limited to the clinical sample. The last section reports on exploratory analyses which include comparisons of results by sample, and findings related to PRR ratings. Tables are organized for ease of comparison, and include both ICS and PRR ratings where applicable.

Descriptive Statistics: Credibility (ICS) and Personal Reactions (PRR) Ratings

Table 3 shows the mean total scores for ICS and PRR ratings by sample. Higher scores on the ICS measure reflect higher credibility ratings, considering dimensions of being logical, scientific, helpful to all depressed people, helpful in other domains of a person’s life, and likely to be effective. Higher scores on the PRR reflect higher personal evaluation of the intervention, related to whether it would be personally regarded as helpful, useful in understanding the cause of depression, likely to improve coping, and effectively treat depression. Mean scores on the PRR measure are lower than ICS scores, due to fewer items on the former scale.

Table 4 shows the correlations between PRR and ICS ratings of the interventions, by sample, with the effect of viewing order partialled out (for reasons described below). ICS and PRR ratings were highly and significantly correlated in the student sample, both
within and across treatment type (viz., cognitive and behavioral). In the patient sample, vignette ratings were correlated within, but not across type; that is, PRR and ICS ratings of the behavioral intervention were not significantly correlated with PRR and ICS ratings of the cognitive intervention.

A difference score was created for each participant for both ICS and PRR ratings by subtracting the total score for the cognitive vignette from the total score for the behavioral vignette. Therefore, these scores are scaled so that higher values reflect higher relative evaluations of the behavioral intervention. To facilitate interpretation, these difference scores are referred to as “ICS-BvC” and “PRR-BvC” scores, where the suffix “BvC” is included to remind readers that higher scores reflect relatively higher behavioral ratings on the relevant measure. Table 5 shows the mean ICS-BvC and PRR-BvC scores, by sample type, along with the ranges these values take within each sample. In the student sample, the mean of the ICS-BvC scores was slightly positive, and the mean PRR-BvC score was near zero. By contrast, in the patient sample, both the ICS-BvC and PRR-BvC means were negative, reflecting higher ratings for the cognitive intervention.

**Considering viewing order.** Viewing order was investigated as a predictor of ratings of the interventions based on prior research using vignettes (Addis & Carpenter, 2000). In the undergraduate sample, 47% of participants (n = 92) watched the behavioral vignette first. A paired t-test identified a significant difference in vignette ICS ratings by viewing order, such that the first video was rated on average 1.58 points lower, therefore less credible, than the second, \( t(195) = -4.7, p < .0001, d = -.67 \). Despite using the same randomization method as the student sample, more subjects in the patient sample viewed the behavioral vignette first (58%; n = 22).iii A similar presentation order effect was
identified in the patient sample, such that the first video was rated as less credible by an average of 2.3 points than the second video ($t(37) = -2.75, p = .009, d = -.90$).

**Hypothesis 1: Credibility (ICS) Ratings by Intervention Type**

Hypothesis 1 predicted that participants would rate the behavioral and cognitive interventions as comparably credible on the ICS. To evaluate how viewing order might influence ICS ratings, repeated measures ANCOVA was utilized in place of paired t-tests, with ICS as the dependent variables, being predicted by viewing order (a between-subjects factor) and by vignette type (a within-subjects factor).

In the student sample, as expected from the paired t-tests, viewing order was a significant predictor of ICS ratings, ($t(195) = -3.56, p = .0005, d = -.51$) with the previously observed result of lower credibility scores for the vignette that was viewed first. There was no significant interaction between viewing order and vignette type for the ICS ($t(193) = -.40, p = .69, r = -.03$), so viewing order could be retained as a simple covariate in the student sample. That is, there was no evidence that the effect of viewing order was influenced by the specific sequence of the vignettes (e.g., higher scores for the second vignette, but only when cognitive was viewed last). In the patient sample, viewing order significantly predicted ICS ratings ($t(37) = -2.54, p = .016, d = -.84$), with this finding once again denoting lower credibility score for the first vignette. There was no significant interaction between viewing order and vignette type for the ICS ($t(35) = -.16, p = .87, r = -.03$), so viewing order could be retained as a simple covariate in the patient sample.

With viewing order covaried, ICS scores did not differ significantly by vignette type in the student sample ($t(194) = .77, p = .44, r = .06$) or the patient sample, $t(36) = -
In summary, repeated measures ANOVA analyses (in which viewing order was controlled) suggest no differences in ICS ratings of behavioral and cognitive interventions for either the student or patient sample, consistent with Hypothesis 1.

**Hypothesis 2: Relation between RFD Score and ICS Ratings**

Higher RFD scores were hypothesized to predict lower ICS ratings of behavioral interventions both in general (Hypothesis 2A), and relative to ratings of cognitive interventions (Hypothesis 2B). As in Hypothesis 1, viewing order was covaried in these regression analyses. Results of analyses testing these hypotheses in the student and patient samples are presented in Table 6. In the student sample, after covarying viewing order, RFD total score did not predict ICS ratings of behavioral or cognitive vignettes, nor did it predict the difference between these two scores (viz., ICS-BvC). In the patient sample, RFD total score was not significantly predictive of ICS ratings of the behavioral vignette. However, RFD total score predicted ICS ratings of the cognitive vignette at a trend level \( p = .06, r = .31 \), such that higher reason-giving was associated with higher credibility ratings of this vignette. Consistent with Hypothesis 2B, RFD total score was significantly predictive of the ICS-BvC \( p = .04, r = -.35 \), showing an association between higher RFD scores and greater relative credibility of cognitive over behavioral approaches.

In summary, there was partial support for Hypothesis 2, but only in the patient sample, and in a somewhat unexpected fashion. There was no evidence supporting either hypothesis in the student sample. By contrast, in the patient sample, RFD scores were predictive of higher credibility ratings of the cognitive intervention at a trend level, and
significantly predicted higher credibility ratings of the cognitive treatment relative to the behavioral intervention. Thus, the predicted relationship between RFD and relatively lower credibility rating of the behavioral intervention was confirmed, but RFD was not significantly predictive of lower ratings of the behavioral intervention in general.

**Hypothesis 3: RFD and Credibility Ratings in the Prediction of Treatment Preference Ratings.**

Hypothesis 3 was that both RFD and ICS-BvC scores would predict ratings of treatment preference. As noted previously, treatment preference ratings were scaled so negative scores reflected cognitive preference, and positive scores reflected behavioral preference, with zero as a neutral midpoint.

In the student sample, both the median and mean were zero (SD = 1.7). The neutral option was selected by 15% of the sample (n=30). The two most common responses were a moderate preference for behavioral (n =43; 22%), and a moderate preference for cognitive (n=40; 20%). Approximately one-third of subjects indicated that they would choose another option if it were available in favor of either cognitive or behavioral approaches (n=63; 32%, with 2 missing data). When participants who disliked both psychotherapies were removed, the mean score was similar at -.12 (SD = 1.7; n=133) and the median remained a zero. These participants overwhelmingly selected the neutral preference option. The patient group had a mean preference score of -.57 (SD = 1.7) and median of 0. The neutral option was the most commonly selected response (n=9; 24%). Only one subject indicated a preference for another option if it were available over either cognitive or behavioral approaches. Viewing order significantly predicted preference ratings in the student sample ($t(194) = -2.20, p = .03, r = -.16$) and
trended toward predicting in the patient group ($t(36) = -2.02, p = .050, r = -.32$), with both findings suggesting greater preference for the last viewed vignette. Therefore, viewing order was covaried in all models.

Before conducting analyses relevant to Hypothesis 3, correlations between potential predictors (i.e., RFD scores and ICS-BvC) were examined in both the student and patient sample, with viewing order partialled out. RFD score was not significantly correlated with ICS-BvC in the student sample ($r = -.05, p = .52$); however, RFD and ICS-BvC were significantly correlated in the patient sample, $r = -.35, p = .04$. This association reflects a relationship between higher RFD scores and relatively higher ICS scores for the cognitive intervention.

Table 7 depicts the results of analyses of predictors of preference in the student sample. ICS-BvC score was robustly predictive of preference ratings, when evaluated individually ($p < .0001, r = .57$), and also when included in a model with RFD total score, $p < .0001, r = .57$. This pattern of results confirmed the expected association between ratings of credibility and preference: higher credibility scores for a given intervention were predictive of preference for that intervention. RFD total score was a trend level predictor of preference, both on its own ($p = .07, r = -.13$) and when ICS difference score was also included in the model ($p = .08, r = -.13$). Higher RFD scores were associated with greater cognitive preference (corresponding to more negative preference ratings). When students who indicated they would choose a treatment other than cognitive or behavioral (n=63) were removed, results were very similar, except that RFD was no longer a trend-level predictor of preference when included with ICS difference score, $t(129) = -1.51, p = .13, r = -.13$. 

52
Table 8 depicts the results of analyses of predictors of preference in the patient sample. ICS-BvC was robustly predictive of preference ratings, as a separate predictor \( p < .0001, r = .78 \) and also when included in a model with RFD total score, \( p < .0001, r = .74 \). As in the student sample, credibility scores were predictive of preference such that higher credibility ratings for a given intervention were predictive of having a preference for that approach. RFD total score was also a predictor of preference on its own \( p = .02, r = -.40 \), but not when ICS-BvC was included in the model, \( p = .46, r = -.13 \). Higher RFD scores were associated with lower preference ratings, corresponding to the predicted preference for cognitive interventions.

**Hypothesis 4: Prediction of Treatment Outcomes from RFD scores, Credibility Ratings, and Preference.**

Hypothesis 4 examined ICS-BvC scores, RFD scores, and preference ratings as predictors of treatment outcomes (viz., symptom change and dropout). All analyses related to Hypothesis 4 were conducted in the patient sample. Analyses involving symptom change consider scores on the Hamilton Rating Scale for Depression (HRSD), which was administered at the last intake appointment, again approximately four weeks later, and again eight weeks after intake. The intake assessment was completed by 38 patients, 32 completed the week 4 HRSD and 29 completed the week 8 HRSD. Table 9 shows HRSD scores and the timing of assessments, overall and by treatment condition, for intake, week 4 and week 8 assessments. Results in the Tables section are organized by analysis type (i.e., symptom change or dropout), but for the sake of clarity, hypotheses are addressed in order below.
**Treatment match variable.** For Hypothesis 4, all analyses involving preference ratings were conducted primarily with the continuous rating data described in Hypothesis 3. However, as prior studies have more often used dichotomous or categorical preference ratings (Zoellner et al., 2003; Zoellner et al., 2009), a treatment match variable was constructed for secondary analyses with categorical preference data. As reflected in the continuous preference scores, patients generally preferred the cognitive intervention. Nineteen patients preferred cognitive treatment (50%), ten preferred behavioral (26%) and nine selected the neutral option (24%). Participants’ preference responses were categorized relative to the treatment they received, with the assumption that patients with neutral preferences would be equally accepting of either treatment and were thus “matched” regardless of what treatment they received. Table 1 provides descriptive information on categorical preference ratings and match variables, as well as showing this data by treatment assignments. 17 patients were mismatched (45%) and 21 were matched or neutral (55%).

**Demographic, diagnostic, and treatment characteristics by intervention assignment.** Table 11 provides information on patient demographic, diagnostic, and treatment characteristics, in addition to tests of between-condition (i.e., behavioral or cognitive intervention) differences. There were no significant differences between interventions on demographic characteristics, including age, sex, race (binary coded as 1 for non-Caucasian participants), or relationship status (binary coded with 1 for “currently single”). Similarly, the intervention groups did not differ in terms of diagnostic characteristics, which included number of prior episodes of depression, diagnosis of chronic depression and/or dysthymia, diagnosis of at least one comorbid anxiety disorder,
or diagnosis of at least one comorbid personality disorder. The interventions also did not differ on the average number of sessions, or patients’ use of concurrent anti-depressant treatment.

**Symptom change analyses.** A longitudinal random coefficients analysis was conducted using PROC Mixed. For these analyses, a vector of HRSD scores was constructed for each patient that included all available HRSD scores from intake, week 4 and week 8 assessments. In this model, the intercept and slope of scores over time were treated as random effects. A significant interaction between a predictor variable and the time variable denotes the prediction of slope of symptom change; in-text descriptions of predictors refer to this interaction effect. A three-way interaction between the predictor, time, and treatment assignment indicates additionally that slopes may differ by treatment condition. Several covariance structures were compared (viz., Toeplitz, compound symmetry, unstructured), with an auto-regressive structure providing the best fit (based on -2 res log likelihood, Akaike’s information criterion, and Schwarz’ Bayesian criterion).

Prior to conducting analyses connected to Hypothesis 4, several treatment variables were considered as potential predictors of symptom change and possibly inclusion as covariates. The effect of time was significant, with the direction of effect suggesting that HRSD scores were decreasing over time, $t(31) = -3.85, p = .0006, r = -.57$. The interaction of the treatment assignment and time was not significant, indicating that interventions did not differ on slope of symptom change, $t(29) = .09, p = .93, r = .02$. Therapist did not predict slope of symptom change, $F(5,29) = .45, p = .96$. Concurrent treatment with anti-depressant medication significantly interacted with time
(t(28) = -2.19, p = .04, r = -.38), with the direction of effect and visual inspection of the
data indicating that patients receiving medication had steeper slopes (more rapid change) than those who did not receive medication. Therefore, this interaction term was retained as a covariate in all analyses from Hypothesis 4.

Table 12 shows results of symptom change analyses related to Hypothesis 4. Two main values are reported for each potential predictor: its interaction with the time variable (a test of prediction of slope of symptom change), and the three-way interaction between the variable, time, and treatment assignment (a test of differential slopes on the basis of treatment assignment). As viewing order was previously established as a predictor of vignette ratings, standardized residuals outputted from a regression of viewing order on ICS difference scores were used as predictors in the symptom change analyses.

**Dropout analyses.** All analyses related to dropout focus on the period of treatment from intake to the week 8 assessment. Nine of 38 patients dropped out before completing the week 8 assessment (24%); all of these patients had at least one therapy session before dropping out. Five of these patients dropped from the cognitive condition. Logistic regression analyses were performed using SAS Proc LOGISTIC, with continuous predictor variables standardized to simplify interpretation. HRSD scores at intake were included as a covariate in all models in order to control for differences in symptom severity that might impact dropout rates. However, intake HRSD did not predict dropout (OR = .84, χ² = .21, p = .65), nor did it interact with treatment to predict dropout (OR = .59, χ² = 1.67, p = .20). Treatment assignment was not predictive of odds of dropout (OR = 1.3, χ² = .15, p = .7). Therapist did not predict dropout (χ² = 1.25, p = .94).
Table 13 provides information on patient demographic, diagnostic, and treatment characteristics by dropout status, and tests of differences between groups (i.e., dropout/non-dropout) on these variables. There were no significant differences between dropouts and non-dropouts on any of these characteristics; therefore, no additional demographic, diagnostic or treatment covariates were retained in the logistic regression models. Table 14 shows the results of the dropout analyses for intervention ratings, preference scores, and RFD; these findings are reported in further detail below by each sub-hypothesis. As in the symptom change analyses, standardized residuals of ICS-BvC scores were used as predictors in the dropout analyses to account for the potential influence of viewing order.

**Hypothesis 4A: RFD.** Hypothesis 4A was that the interaction of reason-giving tendency (i.e., high RFD scores) and treatment assignment would predict both symptom change and odds of dropout. The expected direction of this effect would result in worse outcomes for those high in reason-giving in the behavioral condition.

As shown in Table 12, concerning symptom change across the first eight weeks of treatment, the interaction of RFD total score and time did not predict HRSD slope \( p = .23, r = .23 \), nor did the three-way interaction between RFD total score, time, and treatment assignment, \( p = .39, r = .16 \). Thus, there was no evidence that RFD score was predictive of symptom change, overall or as a function of treatment assignment.

In the dropout analyses, RFD total score predicted dropout risk at a trend level (OR = 2.4, \( \chi^2 = 3.2, p = .07 \)), with the direction of effect suggesting that higher RFD scores were associated with greater risk of dropout across both conditions. However, the interaction of RFD score and treatment assignment was also significant, OR = .20, \( \chi^2 = \)
4.5, \( p = .03 \). One method of interpreting interaction effects in logistic regression involves plotting the predicted probability of dropping out for several fixed scores of the continuous variable (RFD score) at each level of the categorical variable (treatment), holding the continuous covariate intake HRSD at its mean level, with the effect of medication status covaried (UCLA Statistics Consulting Group, 2013).

Regression parameters were outputted from the logistic regression analysis described above. Next, a dataset was created with all possible combinations of predictor variable values of interest (i.e., values of RFD at each level of treatment). Finally, the regression parameters were used in an analysis with the generated dataset, yielding predicted probabilities, which were then graphed. Figure 1 depicts the results of this analysis, showing a modest decrease in probability of dropout in the behavioral condition as RFD increases, and considerably increased probability of dropout in the cognitive condition as RFD increases. Indeed, RFD significantly predicted dropout in the cognitive condition (OR = 18.4, \( \chi^2 = 3.9, p = .049 \)) such that higher RFD scores predicted greater odds of dropout; no relationship was observed in the behavioral condition, OR = .88, \( \chi^2 = .05, p = .83 \). This was contrary to the expected pattern of association between RFD and dropout.

In summary, there was no support for Hypothesis 4A on the basis of these analyses. RFD total scores did not predict symptom change, either alone or interacting with treatment assignment. RFD total score was predictive of dropout odds in general, with higher RFD associated with higher odds of dropout, though this finding was not predicted. Paradoxically, the interaction of RFD with treatment assignment was significant but in an unexpected fashion, such that higher RFD scores were predictive of
greater probability of dropout in the cognitive treatment, and somewhat associated with lower probability of dropout in the behavioral treatment. Thus, the identified relationship between RFD and dropout was not as hypothesized.

**Hypothesis 4B: Credibility.** Hypothesis 4B was that the difference between credibility ratings of the two interventions (viz., ICS-BvC) would be predictive of symptom change and dropout. The expected direction of this effect was that credibility scores more favorable to the assigned treatment would be associated with superior outcomes (i.e., more rapid symptom change, lower odds of dropout), meaning that higher ICS-BvC scores should predict better outcomes in the behavioral condition, evident by an interaction with treatment assignment.

As shown in Table 12, the interaction of time and ICS-BvC score was not predictive of slope of symptom change, $p = .93, r = .02$. The three-way interaction of treatment assignment, ICS-BvC score, and time also failed to predict symptom change, $p = .84, r = .04$. In the logistic regression model (Table 14), ICS difference scores were not predictive of dropout (OR = .60, $\chi^2 = 1.6, p = .20$), and the interaction between ICS-BvC score and treatment assignment also failed to predict dropout, OR = 1.41, $\chi^2 = .69, p = .41$. In summary, Hypothesis 4B was not supported by the data, as ICS-BvC scores failed to predict symptom change or odds of dropout.

**Hypothesis 4C: Preference and treatment match.** Hypothesis 4C was that patients’ preference ratings would predict symptom change and dropout odds, such that patients receiving a preferred intervention would evidence better outcomes (i.e., steeper slope of symptom change; lowered odds of dropout). To evaluate this claim, preference ratings were considered both as a continuous variable (viz., ratings of strength of
preference, including the neutral option), and as a categorical variable (viz., the matching variable described above).

Table 12 shows results of the symptom change analyses. Continuous ratings of preference did not predict HRSD slope overall \( (p = .74, r = .06) \), nor did preference rating interact with treatment assignment to predict symptom change, \( p = .82, r = .05 \). Treatment match did not predict symptom change \( (p = .72, r = -.07) \), and the interaction of treatment match and treatment assignment also failed to predict symptom change, \( p = .25, r = -.22 \).

Concerning dropout, the continuous treatment preference variable was a significant predictor of dropout odds \( (OR = .49, \chi^2=5.1, p = .02) \), such that higher preference ratings (reflecting a relative preference for behavioral treatment) were associated with lower odds of dropout. Surprisingly, the interaction between treatment assignment and preference ratings did not significantly predict dropout. This suggests that simply having a preference for behavioral treatment was associated with lower odds of dropout, and that this effect was not strongly moderated by the actual treatment to which a client was assigned. The categorical treatment match variable was not predictive of dropout odds \( (OR = 1.7, \chi^2 = .40, p = .52) \), nor did the interaction of treatment match and treatment type predict dropout odds \( (OR = .71, \chi^2 = 1.8, p = .17) \).

In summary, Hypothesis 4C received little support. There was no relationship between treatment preference and symptom change, but some evidence of a relationship between preference ratings and dropout. However, this relationship was not consistent with hypothesized relations, as greater behavioral preference was associated with lower
dropout rates across conditions, and no relationship was detected between treatment match/mismatch and dropout.

**Exploratory Analyses: Comparing Significant Predictors of Dropout**

In the service of a complete comparison, RFD total score and continuous treatment preference ratings were compared in the prediction of dropout, as both of these variables predicted dropout. In this dataset, these two variables were not significantly correlated ($r = -.28, p = .09$) but trended toward a negative relationship, reflecting the expected association between cognitive preference and higher RFD scores. When both predictors were added to the logistic regression, after controlling for concurrent medication status, RFD was no longer predictive ($OR = 2.2, \chi^2 = 1.8, p = .18$), while preference ratings predicted at a trend level ($OR = .53, \chi^2 = 3.8, p = .05$).

**Exploratory Analyses: Post-hoc Comparison of Students and Patients**

Exploratory analyses were conducted to examine whether the associations tested in the first three hypotheses differed as a function of sample. In these analyses, an interaction term representing a group difference (i.e., student versus patient sample) was tested. Viewing order was covaried in all analyses. In terms of the predictor variables of interest, patients tended to give higher credibility scores to both interventions ($M = 25.1, SD = 4.1$) compared to students ($M = 23.7, SD = 4.5, t (232) = -2.59, p = .01, d = -.34$). ICS-BvC scores were not predicted by sample ($t (231) = 1.64, p = .10, r = .11$). As reported earlier, patients had considerably higher RFD scores ($M = 23.1$ for patients, $M = 11.3$ for students). Patients and students also differed significantly in terms of preference ratings ($t(232) = 2.03, p = .04, d = .27$), with the mean patient score being negative ($M = -.58, SD = 1.7$) and therefore reflective of a tendency to prefer cognitive treatment, versus
a near-zero score for the student sample reflecting no clear group preference (M = .03, SD = 1.7).

For Hypothesis 1, the interaction of intervention (i.e., cognitive or behavioral) and sample (i.e., student or patient) was not significant for ICS ratings ($t(231) = 1.28, p = .20, r = .09$), indicating that the two samples did not differ on their credibility ratings of the interventions. For Hypothesis 2, the interaction of sample and RFD score was not significant in predicting either ICS ratings of either the behavioral intervention ($p = .60$) or cognitive intervention ($p = .20$). This suggests that the relationship between RFD score and ratings of each individual intervention did not differ as a function of sample.

Similarly, the interaction of RFD and sample was not significant in predicting of ICS-BvC score, $t(229) = -1.66, p = .10, r = -.11$. Finally, for Hypothesis 3 in the combined sample, there was no interaction between RFD and sample in the prediction of preference ratings, $t(229) = 1.32, p = .19, r = .09$. The interaction of ICS-BvC score and sample also failed to significantly predict preference, $t(229) = -1.25, p = .21, r = -.08$. The absence of any significant interactions across the main hypotheses suggests that sample did not significantly moderate the relationships of interest.

**Exploratory Analyses: Personal Reactions to the Rationales (PRR) Ratings.**

This section describes all analyses involving PRR ratings, which were included in the study but not incorporated into primary hypotheses due to this measure being less well-established and connected to the relevant literature than the credibility scale. The PRR scale differs from the ICS scale primarily in terms of the focus on personal evaluation and choices. PRR ratings were examined as predictors in a series of analyses that paralleled those involving ICS ratings, including all relevant covariates. Note that
results from this section are presented in tables alongside previously reported findings with ICS ratings.

**PRR Ratings by Intervention.** In parallel to Hypothesis 1, differences in PRR ratings by vignette were examined in both the student and patient sample, after covarying viewing order. PRR scores did not differ by intervention in the student sample \((t(194) = -.12, p = .91, d = -.02)\), but were different in the patient sample \((t(36) = -3.02, p = .0046, d = -1.0)\), with patients generally giving higher PRR ratings for the cognitive vignette \((M = 18.6, SD = 3.8)\) versus the behavioral vignette \((M = 15.9, SD = 3.4)\). This contrasts with findings from the main study hypotheses, as neither student nor patient samples evidenced a difference between interventions on credibility ratings.

**Relationship between PRR Ratings and RFD.** Next, RFD scores were examined as predictors of both the individual intervention PRR ratings (Hypothesis 2A) and the PRR-BvC score (Hypothesis 2B). Results are shown in Table 6. In the student sample, RFD total score predicted behavioral PRR rating \((p = .03, r = -.15)\), with higher RFD score predicting lower ratings of behavioral vignettes. RFD total score did not predict cognitive PRR ratings, or PRR difference score. This finding differs from the main study analyses in the student sample, in which RFD failed to predict ICS scores for either intervention or ICS-BvC scores. However, this result is in keeping with the main hypothesized association of higher RFD scores predicting less favorable evaluation of behavioral treatment.

In the patient sample, RFD total score failed to predict scores of the behavioral vignette, but was significantly predictive of both the cognitive vignette PRR rating \((p = .01, r = .40)\) and the PRR-BvC score \((p = .01, r = -.41)\), with the direction of the effect
in both cases showing higher RFD associated with more positive PRR ratings of the cognitive intervention. This pattern of results parallels findings with the ICS in the patient sample, with RFD scores predicting both cognitive ICS and ICS-BvC scores, such that higher RFD scores predicted more favorable ratings of the cognitive intervention.

**PRR Ratings and Treatment Preferences.** Hypothesis 3 investigated predictors of preference, including RFD scores and rating difference scores (viz., ICS-BvC and PRR-BvC). PRR-BvC score was tested as a predictor of preference ratings, alone and in a model that also included RFD ratings. RFD and PRR-BvC were not significantly correlated in the student sample \( r = -0.09, p = .19 \), but were significantly correlated in the patient sample, \( r = -0.44, p = .01 \). Table 7 shows the results of these analyses in the dysphoric undergraduate sample. PRR-BvC was robustly predictive of preference ratings, as an individual predictor \( p < .0001, r = .62 \), and also when included in a model with RFD total score, \( p < .0001, r = .61 \). Table 8 shows results from analyses of predictors of preference in the patient sample. As in the student sample, PRR-BvC predicted preference both as an individual predictor \( p < .0001, r = .80 \), and when included in a model with RFD total score, \( p < .0001, r = .76 \). This pattern of results was consistent with findings from the ICS and study hypotheses, such that higher relative rating on the PRR for a given intervention predicted preference for that intervention.

**PRR Ratings and Treatment Outcomes.** Hypothesis 4B concerned the relationship between intervention ratings and two treatment outcomes: early symptom change, and dropout. All treatment outcome analyses were conducted in the patient population, after controlling for concurrent anti-depressant treatment status. The relationship between PRR-BvC and therapy outcomes was tested as a main effect, as well
as in interaction with treatment type (i.e., cognitive or behavioral). As reported in Table 12, PRR-BvC failed to predict early symptom change as a separate predictor \((p = .65, r = .09)\), nor did the interaction of PRR-BvC and treatment assignment predict early symptom change, \(p = .54, r = .12\). This result mirrors the failure to detect a significant predictive relationship between ICS-BvC score and symptom change, either overall or the interaction of that score and treatment assignment.

As described in Hypothesis 4, in dropout analyses, intake HRSD score was included as a covariate, along with a binary variable indicating whether patients were receiving medication. As reported in Table 14, PRR-BvC score did not predict dropout odds overall \((OR = .84, \chi^2 = .18, p = .67)\) but the interaction between PRR-BvC and treatment assignment was predictive at a trend level, \(OR = 2.2, \chi^2 = 3.2, p = .07\). As in Hypothesis 4A, interaction effects were interpreted by plotting values of the continuous variable (standardized PRR-BvC) at each level of the categorical variable (treatment), covarying medication status and holding intake HRSD at its mean. As shown in Figure 2, changes in PRR-BvC scores favoring the assigned treatment were associated with higher probability of dropout in that treatment, with this effect more apparent in the cognitive treatment. However, PRR-BvC score was not a statistically significant predictor of dropout in either the cognitive condition \((OR = .44, \chi^2 = 1.5, p = .22)\) or behavioral condition, \(OR = 1.8, \chi^2 = .73, p = .39\). Although there were no specific hypotheses related to the PRR, this finding is surprising and not consistent with what might be the expected pattern of association between ratings of personal reactions to a particular intervention and likelihood of dropping out of that intervention.
Comparison of Results with ICS and PRR. In summary, results of analyses with PRR scores in the main study hypotheses were overall very similar to those with ICS scores. Notable exceptions include the finding that patients rated the cognitive intervention higher than the behavioral intervention on the PRR, RFD score predicted lower PRR ratings of the behavioral intervention in the student sample, and the finding that relatively higher PRR ratings of a given intervention were associated with greater probability of dropout in that treatment, perhaps particularly in the cognitive treatment.
Chapter 4: Discussion

This study explored relations between ratings of treatment credibility of cognitive and behavioral interventions, beliefs about the cause of one’s depression, and treatment preferences in a large group of dysphoric undergraduates, and a smaller group of depressed outpatients. Analyses in the patient sample investigated whether these ratings predicted symptom change and dropout. Ratings of personal reactions to the treatment rationales were also investigated in a series of analyses that paralleled the main study hypotheses.

In general, findings were more consistent with hypothesized and theoretical associations between these variables for the patient group versus the dysphoric students, though there were no significant differences by sample. Both groups rated the intervention vignettes and treatment rationales as equally credible; the patient sample had more favorable personal reactions to the cognitive treatment. As expected, high RFD scores predicted higher credibility ratings of the cognitive vignette in the patient sample. High RFD scores also predicted preference for cognitive treatment but this effect was small compared to the relations between preference and ICS or PRR ratings. There were no significant predictors of early symptom change in the treatment study. Several predictors of greater dropout odds were identified, but none of these relationships were consistent with hypothesized relationships.
**Intervention Ratings**

After controlling for viewing order, both patients and undergraduates rated the cognitive and behavioral vignettes and treatment rationales as equally credible. This was consistent with expectations, based on the similarities between the interventions as described in this study, and previous studies involving separate cognitive and behavioral treatments (e.g., Rokke et al., 1990; Rokke & Scogin, 1995). While not designated as a specific test of the suitability of the newly developed vignettes and treatment rationales, the absence of overall differences in credibility ratings (in conjunction with mean scores in the mid-range of the credibility scale) does suggest that the vignettes and rationales are adequate for use in this context.

Supplemental analyses involving the PRR revealed a significant difference between the two interventions in the patient sample. The PRR measure is similar to the ICS but incorporates a more personalized perspective on treatment. Thus, in general, patients seemed to regard the interventions as equally credible, but to hold the cognitive interventions in higher regard concerning their own care, a finding consistent with subsequent assessments of preference. No such differences on PRR ratings were identified in the student group. This may reflect a fundamental distinction between the student and patient sample; namely, that the patients were actively seeking treatment for substantial depressive symptoms, whereas the students as a group tended not to be doing so. Thus, the PRR measure may be assessing some aspect of the attitudes of patients towards their own treatment, which may not have been as salient or well developed in the non-treatment seeking and generally less symptomatic student group. This idea is developed more fully in the section that follows.
Reasons for Depression and Reason-giving

The hypothesized association between treatment ratings and RFD score was based on the theory that individuals with high RFD scores have a tendency to ascribe complex causal factors to their own depression, dubbed reason-giving per Addis and Jacobson (1996), and would therefore regard behavioral interventions as being too straightforward to address their depression. This association between RFD scores and ratings of interventions was evident in this study, though results were not entirely as expected. In general, findings were more consistent in the patient sample, but no direct associations were found between RFD score and either ICS or PRR ratings of the behavioral intervention. Instead, higher RFD scores predicted higher ratings of the cognitive intervention higher on ICS and especially on the PRR, with a corresponding association between RFD score and difference scores on both rating measures. Therefore, in the patient sample, the association between RFD and intervention ratings was consistent with Addis and Jacobson’s reason-giving theory, such that individuals with higher RFD scores tended to regard the more complex intervention as more credible and personally palatable.

In the student sample, the expected pattern of association between RFD and intervention ratings was evident only for PRR ratings of the behavioral intervention, and did not extend to corresponding associations with either the cognitive intervention or PRR difference scores, or any ICS ratings. The failure to detect the hypothesized patterns of association in the comparatively larger dysphoric student sample suggests the possibility of meaningful differences in responses between the two samples. Consistent with the theoretical development and application of the RFD inventory, depressed
patients’ ratings were higher than the dysphoric student group, reflecting endorsement of a greater number of possible causal explanations for their depression. The student group also had lower depressive symptoms (i.e., BDI-II scores), and on average rated both interventions as less credible and personally appealing than the patient group.

It is possible that the student sample’s responses on the RFD were in some way qualitatively different than the patient sample. Though previous studies have used the RFD in non-clinical populations (e.g., Addis & Carpenter, 1999), the primary aim of the measure is to distinguish individuals’ beliefs about the causes of their own depression, as opposed to beliefs about how depression could be caused in general (Addis et al., 1995). Addis and Jacobson (1996) imply that it is the personal component of reason-giving that influences depressed persons’ reactions to treatment options, as opposed to global perceptions of the multiple factors that might cause anyone’s depression. Absent the perception of currently being depressed, RFD scores may no longer be a reflection of the personal component that is central to the reason-giving style. Therefore, in the student sample, only the subset of participants who regard themselves as being depressed, and whose responses to the RFD are influenced by that perception, would be responding in a fashion similar to the original reason-giving hypothesis put forth by Addis and Jacobson (1996), and more similarly to the patient sample in this study, all of whom were seeking treatment for depression.

Exploring the relationships between RFD, symptoms of depression, and intervention ratings in both participant groups yields additional information that suggests meaningful differences in patterns of responding between the two samples. Addis and Carpenter (1999) reported a significant correlation of $r = .76$ between RFD and BDI-II
score in their non-clinically depressed sample. In this study, BDI-II scores and RFD were significantly correlated in both students ($r = .57$) and patients, $r = .60$. In the patient sample, BDI-II scores significantly predicted PRR behavioral ($t(32) = -2.9$, $p = .008$, $r = -.45$), cognitive ($t(32) = 3.2$, $p = .003$, $r = .50$), and PRR-BvC scores ($t(32) = -4.5$, $p < .0001$, $r = -.60$), with increased symptoms associated with more favorable ratings of the cognitive treatment and less favorable ratings of the behavioral treatment, separately and relative to one another. The pattern of results observed in the patient group is consistent with the pattern of results observed by Addis and Carpenter (1999): higher RFD scores correlated with higher symptoms and predicting less favorable ratings of the behavioral treatment. However, in the student sample, BDI-II scores significantly predicted both behavioral ($t(193) = -3.5$, $p = .0006$, $r = -.24$) and cognitive ($t(193) = -2.2$, $p = .03$, $r = -.15$) PRR ratings, but in both cases higher levels of symptoms were associated with less favorable PRR scores. Furthermore, BDI-II scores did not significantly predict PRR-BvC score in the student sample, $t(193) = -1.05$, $p = .30$, $r = -.08$. This finding suggests important differences between the two samples, indicating that the dysphoric undergraduates’ responses were not in line with previously observed relationships between depressive symptoms, RFD, and ratings of the two interventions that were evident in the patient sample.

**Preference Ratings**

The patient and student samples differed significantly in terms of treatment preference ratings ($d = .27$), with the average score in the patient group reflecting a slight preference for cognitive treatment, versus a neutral preference in the student sample. Approximately one-third of the student group indicated that they would prefer a treatment
other than the cognitive or behavioral interventions being depicted in the vignette, versus only a single participant in the treatment sample. While this suggests another possibly meaningful conceptual difference between the two groups, the pattern of findings with respect to prediction of preference was quite similar in both samples.

In the student sample, RFD scores were predictive of preference ratings at a trend level, such that higher RFD scores were associated with greater cognitive preference. However, both ICS-BvC and PRR-BvC scores were more robust predictors of preference ratings than RFD score. Results were similar in the patient sample, except that the association between RFD and preference was significant when tested as a separate predictor, and non-significant when included with either ICS-BvC or PRR-BvC scores. In post-hoc analyses in which both samples were combined, RFD was significantly related to preference as a sole predictor, a trend level predictor when included with ICS-BvC score, and no longer significant if PRR-BvC score was included in the model.

In summary, the associations between RFD, intervention ratings, and preference were broadly consistent across the first three sets of hypotheses, even in spite of differences between the two samples. Results of analyses involving ICS-BvC and PRR-BvC scores were similar. In the patient sample, RFD scores showed the anticipated pattern of association with intervention ratings (i.e., higher RFD reflecting better ratings of the cognitive treatment). Though RFD was associated with preference ratings as hypothesized, its ability to predict preference ratings was small in comparison to ICS-BvC and PRR-BvC scores.
Prediction of Outcomes from Cognitive and Behavioral Intervention Pilot Study

Results from the first eight weeks of the treatment study suggest that both intervention groups showed a reduction in symptoms. Dropout rates were also comparable across treatment conditions. As this study concerns only an eight-week treatment interval, comparison of these outcomes to the published literature are somewhat difficult, as studies predominantly involve treatments of longer or non-fixed intervals. Three studies (DeRubeis et al., 2005; Watkins et al., 1993; Wright et al., 2005) provide useful comparisons, as they involve CT conditions, and reporting of outcomes with the 17-item HRSD at an eight-week interval.

DeRubeis and colleagues’ (2005) large trial of CT versus medication may be least ideal as a comparison, as all patients were required to have HRSD $\geq 20$ at intake ($M = 23.4$, $SD = 2.9$ across all conditions, versus $M = 20.2$, $SD = 5.2$ in the present study) and results are reported only in terms of response percentage. 43% of the patients in the intent-to-treat CT condition (n=60) were considered responders at week 8, based on having HRSD scores $\leq 12$. In the current study, 40% of patients (15 of 38) were considered responders by week eight, based on the same HRSD criteria and last-observation-carried-forward method for dropouts.

Watkins and his colleagues (1993) reported on the CT condition of the Treatment of Depression Collaborative Research Program (TDCRP), with all patients required to have HRSD scores $\geq 14$ at intake. Their sample of treatment completers (n=36) had a mean HRSD of 19.1 ($SD = 4.2$) at intake, and 11.9 ($SD = 6.5$) at week 8, with a pre-post effect size of $d = 1.7$. The comparable sample from the present study (n=27) had a mean HRSD of 20.7 ($SD = 4$) at intake and 14.3 ($SD = 5.6$) at week 8, with $d = 1.6$. 
Wright and colleagues’ (2005) evaluation of computer-assisted therapy also included a traditional CT condition (n=15), without an intake HRSD severity requirement. The mean intake HRSD in their CT condition was 17.1 (SD = 5.4), with intent-to-treatment outcomes reported as a mean improvement of 8.3 points (SD = 6.6), equivalent to an unbiased $d$ of 1.2 (Schinka & Velicer, 2013, pp 159). In the present study, mean HRSD change from intake to week eight was 5 (SD = 6) for the intent-to-treat group, equivalent to an unbiased $d$ of .85. Completer analyses yield an unbiased $d = 1.3$ for the current study versus $d = 1.2$ for the Wright et al. (2005) CT condition. In both sets of analyses, effects of treatment were large, with intent-to-treat analyses showing the largest discrepancy in effect size across studies. Thus, effect sizes of symptom change and rate of responding in this 8-week trial were generally comparable to other trials that reported outcomes across a similar time period.

Considering dropout rates, the 24% dropout rate observed in the first eight weeks of this study is higher than both the 15% reported by DeRubeis and colleagues (2005) and 13% reported by Wright and colleagues (2005) across the first eight weeks of their respective CT conditions (dropout could not be calculated for the eight week interval of the TDCRP). However, this estimate it is in keeping with the range of recent meta-analytic estimates from the literature across treatments of varied lengths (17-19%; Cooper, 2013). It also bears consideration that dropout rates appear to be higher with trainee therapists (27%) and in university based clinics (30%) than other settings (Swift & Greenberg, 2012).

**Symptom change.** There were no significant predictors of symptom change across the first eight weeks among the rated variables (i.e., ICS or PRR difference scores,
preference ratings, or RFD scores) or the treatment match variable. In the context of this pilot study of cognitive and behavioral interventions delivered separately in the treatment of depression, patients’ evaluations of the treatments being offered, their treatment preferences, and whether or not they received their preferred treatment, did not substantially predict their rate of symptom improvement. Given the overall tendency for patients to improve over the course of this time period, this finding suggests the possibility that these variables may not be especially informative in understanding what predicts patients’ symptom improvements. This is particularly surprising given the finding that RFD was predictive of treatment response within the behavioral condition of the dismantling study (Addis & Jacobson, 1996). The most straightforward interpretation of this result is that that patients’ preferences (and other related variables) do not dramatically influence the rate of response to treatment. However, there are a number of additional factors worth considering before accepting this interpretation.

First, the majority of prior research on the relation between patient preferences and treatment outcomes (including symptom change) has involved comparison of overtly different treatment modalities, most commonly pharmaceutical versus psychotherapy interventions (e.g., Zoellner et al., 2003). Preference effects would likely be more pronounced between interventions of these kinds. By contrast, this study compares two interventions that are typically subsumed in one overall treatment approach, setting it apart even from prior research comparing theoretically distinct forms of psychotherapy (e.g., Rokke et al., 1990; Addis & Carpenter, 1999). Judgments of treatment credibility might be expected to be comparatively less discrepant, and variability in preferences might be less substantial than studies involving interventions with more overt differences.
These same considerations apply with respect to the failure to identify any association between treatment matching and symptom change. Due to patients being more likely to prefer the cognitive treatment, there were higher rates of preference mismatch between preferred and received treatment in the behavioral condition. As this study did not involve randomization on the basis of treatment preference (in either the preference or PRPT formats discussed earlier), the effect of treatment choice cannot be evaluated. As noted previously, studies involving treatment choice tend to report more robust findings (e.g., Rokke et al., 1999; Swift & Callahan, 2009). However, the substantial number of neutral preference responses (n=9; 24%) further support the contention that preferences – and, by extension, treatment matching – may not be as relevant in predicting symptom change as hypothesized in this study. Considering that all patients were voluntarily recruited to a psychotherapy training clinic suggests that this sample might differ from one in which participants were informed that they might receive a wider range of treatment options (e.g., pharmacotherapy or more dissimilar psychotherapy treatments). Data pertaining to the second half of treatment were not the focus of this project, so it remains to be seen whether effects will be detected in the longer treatment duration (i.e., 16 weeks), or after patients have been randomized a second time.

Considering the surprising failure to identify any association between RFD ratings and symptom change, differences between this study and Addis and Jacobson’s (1996) follow-up to the dismantling study warrant consideration. These include the shorter duration of this study (eight weeks versus 16 weeks), the considerably smaller sample (n=38 versus n=98), the method of describing the intervention (written rationale and
video vignette versus a modified *Coping With Depression* pamphlet), novice versus experienced therapists, and differences in the RFD itself, with the abbreviated version used in the present sample. Any of these differences might have influenced the failure to detect an association. Another notable difference is that the statistical approach utilized in this study is an advancement from the method employed in the earlier study (i.e., HLM analyses versus correlation with a composite HRSD-BDI depression score at post treatment, partialling out pre-treatment score).

By way of comparison, using completer data from Addis and Jacobson’s (1996) study, the correlation between RFD and residualized change in a composite BDI-HRSD symptom measure was statistically significant at $r = .36$ in the BA condition ($n=49$), with the direction of effect showing worse outcomes (i.e., less change) for those with higher RFD scores in the BA condition. In behavioral condition completers from the present study ($n=15$), the correlation between RFD and residualized change in HRSD from intake to week eight was $r = .47$ ($p = .08$), with the same direction of effect. Thus, in comparable analyses, similar relations between RFD and outcome in the behavioral condition were observed. Subsequent analyses in the present dataset will examine RFD and treatment assignment not only across the full 16-week treatment, but also in relation to other outcomes from Addis & Jacobson (1996), including predicting homework compliance, early ratings of treatment helpfulness, and both patient and therapist perceptions of the therapeutic alliance.

**Dropout.** In contrast to analyses focused on symptom change, this study identified several predictors of dropout. The interaction of PRR difference score and treatment assignment trended toward predicting dropout, with the probability of dropping
out of a given intervention increasing as PRR ratings of that intervention were more favorable, particularly in the cognitive condition. Curiously, the trend-level PRR-by-treatment interaction was the only finding related to treatment assignment amongst the multitude of similar and potentially more promising dimensions of preference and treatment matching (i.e., ICS difference scores, preference ratings).

RFD score was also found to predict dropout, both overall and as an interaction with treatment condition, with higher RFD predicting greater odds of dropout in the cognitive condition, and no relationship between RFD and dropout odds in the behavioral treatment. Considering the association between RFD and cognitive preference, and the relatively higher rate of cognitive preference overall, one might expect the RFD-dropout association to reflect greater likelihood of high RFD, cognitive-preferring patients being mismatched to treatment. Unexpectedly, simply being mismatched to treatment was not predictive of higher odds of dropout; in fact, all dropouts in the cognitive condition had been matched to their preferred treatment. Likewise, though RFD was moderately correlated with symptom severity, HRSD score was not predictive of dropout.

The overall predictors of dropout identified in this study (i.e., higher RFD, cognitive preference, and higher PRR ratings of cognitive treatment) all point to an association between favoring the cognitive intervention and dropout, with findings of interaction with treatment assignment (i.e., RFD and PRR-BvC) also supporting this view. Half of the sample (n=19) reported either a neutral or behavioral preference, but only one patient in this group dropped out (5%). The remaining cognitive-preferring patients had a much higher dropout rate (8 out of 19, or 42%). The pattern of results seen in the prediction of dropout suggests the possibility that patients’ may be entering the
study with a characteristic that puts them at greater risk of dropout, over and above their treatment assignment, as the tendency to prefer the cognitive intervention (and its associated measures) was associated with higher dropout rates. Considered another way, patients who favored the behavioral intervention or had neutral preferences were less likely to drop out.

It is possible that some constellation of personality or symptom characteristics is associated with a preference for cognitive treatment and is also a risk factor for dropout. For instance, if patients prefer the cognitive intervention in part due to an accurate perception of struggling with negative thoughts, they may have more difficulty overcoming doubts or negative judgments about remaining in treatment. There might also be some correlate of preference for the behavioral intervention that confers protection against attrition, such as an expectation that treatment may be challenging and require effort on tasks between sessions. A patient entering treatment without a strong preference might be less likely to find an issue with their assigned treatment that was sufficiently motivating to get them to drop out. Considering the RFD-dropout relationship, a patient who does not endorse a complex, multifaceted view of his or her depression (i.e., lower RFD score) may be less likely to have concerns about whether their treatment matches their own view of what may cause or maintain their distress, or whether it is sufficiently complex to adequately treat them.

Limitations

This study has a number of limitations. First, the sample size was smaller than anticipated due to unexpected issues encountered during data collection. The substantial incidence of partial or suspect completion in the student sample suggests a potential
downside to the online-only data collection methods that were employed. An additional complication in the patient group was the unexpected delay in beginning recruitment, and difficulty in recruiting eligible participants, resulting in smaller overall sample size and cells (e.g., treatment match by condition). This affected the ability to detect effects.

Second, this study was intended to connect with a number of different, largely distinct lines of research, including investigations of perceptions of treatment rationales, ratings of preference, and the relationship between beliefs about the cause of depression and options for treatment. The study was designed to maximize comparability across these literatures. However, straightforward comparisons are complicated by differences in method and analyses. Notably, in the most recent study involving most of the main predictor variables of interest (i.e., RFD, ICS, and PRR), Addis and Carpenter (1999) utilized the full version of the RFD, with responses given on a five-point scale (most likely to allow for a neutral response to each item). By contrast, this study utilized the more conventional four-item response scale, as well as employing a shortened version of the form (Leykin et al., 2007). Whereas Addis and Carpenter (1999) had participants rate ICS and PRR separately, more recent studies using the ICS and PRR (e.g., Zoellner et al., 2009) have instructed participants to rate both measures simultaneously and have often combined them, as was done in this study. The aforementioned series of studies conducted by Zoellner, Feeny and their colleagues have utilized both vignette and written rationale methods, separately and together, with participants typically rating both credibility and PRR at the same time. Nevertheless, this method variance complicates direct comparison with the most similar study to date that focuses on treatments for depression (Addis & Carpenter, 1999).
This study also differs from prior research in terms of how preference is assessed. Scalar ratings were chosen to capture dimensions of strength of preference, in the event that patients varied on this dimension. However, in not opting for a binary forced choice set up (akin to the approach used in the PTSD studies of Zoellner, Feeny and colleagues), the resultant data is complicated by the presence of neutral respondents (~25% of the patient sample). Overall, this approach likely preserves meaningful variability, but fundamentally changes the way that participants are allowed to respond, which is a consideration when comparing the results of this study (and the constructed treatment match variable) to other studies. Likewise, this study does not incorporate treatment choice, which has been found to predict more robust preference effects (Swift et al., 2011). However, it is likely that allowing patients to choose their preferred intervention would have resulted in substantially imbalanced treatment groups.

Finally, the use of dysphoric undergraduates as one of the samples in this study warrants consideration. There is a precedent for utilizing non-clinical samples in research of this kind, including many of the prior studies upon which this work is based (e.g., Addis & Carpenter, 1999; Rokke et al., 1990; Zoellner et al., 2003). Furthermore, as shown above, sample did not appear to significantly affect results of the main study hypotheses. However, at least insofar as some of the primary predictors of interest relate directly to an individual’s perceptions of the cause(s) of their depression (i.e., RFD), or their reactions about their own personal treatment (i.e., PRR), the results of this study call into question whether even dysphoric participants are truly comparable to depressed treatment-seeking individuals. Certainly, there is a need to better understand how individuals with some symptoms of depression who are not presently seeking treatment
consider their options and the putative cause of any symptoms they might be experiencing. In fact, this population may be of particular interest to those who are investigating the decision to pursue treatment, or how a typical consumer evaluates treatment options (e.g., Pruitt et al., 2012). Nevertheless, the potential differences in responding between dysphoric undergraduates and currently depressed people who are actively seeking treatment options are worthy of consideration for those who intend to directly compare the two groups. Future research of this kind would benefit from a more heterogeneous community sample to attempt to address some of the differences between these groups.

**Conclusions**

This study aimed to better understand how individuals evaluate the cognitive and behavioral interventions that comprise CBT for depression. Following the model of prior studies that compared different treatments (e.g., medication versus therapy), a novel set of rationales and video vignettes were developed to describe how these interventions are utilized in the treatment of depression. In general, the treatment-seeking depressed patient sample had findings more consistent with hypotheses than the dysphoric undergraduates, though results were similar. Both groups rated the interventions as comparably credible, but the patients had more favorable personal reactions to the cognitive intervention. The pattern of association between intervention ratings and RFD responses seen in the patient group in particular was consistent with Addis and Jacobson’s (1996) reason-giving hypothesis: namely, higher RFD scores were associated with more favorable ratings of the cognitive intervention (in terms of personal reactions, credibility ratings and preferences). Additional analyses to identify predictive relationships between RFD,
interventions ratings, or preferences and treatment outcomes failed to discover any predictors of symptom change. Unexpectedly, predictors of dropout emerged suggesting that patients who rated the cognitive intervention more favorably (in terms of PRR and preference) tended to have higher rates of dropout. Higher RFD scores were also predictive of higher rates of dropout, overall and especially in the cognitive intervention condition.

These results suggest a possible role for RFD and PRR ratings as useful indices of preferences between cognitive and behavioral interventions. These variables were also predictive of odds of dropout, with higher RFD scores and preference for cognitive interventions appearing to be general risk factors for dropout. The paradoxical finding that high RFD, pro-cognitive patients may have been more likely to drop in the cognitive-only intervention employed in this pilot study suggests the need for future investigations into the whether patient retention might actually be improved by using a non-preferred intervention, or by examining how patients with neutral or behavioral preferences might be protected from dropout. Future analyses in this dataset will expand to the full 16-week intervention and consider prediction of other outcomes and process variables, including assessor-rated cognitive and behavioral skill usage, therapeutic alliance, homework compliance, and intervention-specific treatment targets (e.g., dysfunctional attitudes in cognitive treatment; behavioral inactivation in behavioral treatment).
References


* Ward, E., King, M., Lloyd, M., Bower, P., Sibbald, B., Farrelly, S., et al. (2000). Randomized controlled trial of non-directive counseling cognitive-behavior therapy,


The version of the RFD employed by Addis and Carpenter (1999) may have been different than prior or subsequent versions. While it contained the same number of questions as the original long form version (Addis et al., 1995), it is described as a 5-point Likert type scale, possibly reflecting the inclusion of a mid-point or neutral response, in contrast to the typical 4-point response approach, which necessitates choice of a positive or negative response.

Due to scheduling issues, some patients had multiple intake assessments in order to complete all study measures. This second assessment was considered the official start date for these patients, with all subsequent assessments scheduled relative to this date. The HRSD was always completed at the last assessment to ensure the most up to date symptom measurement.

The unequal randomization in the patient sample did not appear to result from the influence of four subjects who were ineligible due to missing data, as they were split evenly between seeing cognitive and behavioral vignettes first.

This participant had previously received CBT treatment.

Firth’s penalized likelihood was used in the test of the interaction of treatment with treatment match variable to address issues with quasi-complete separation of data points.
Appendix A: Tables
<table>
<thead>
<tr>
<th>Study</th>
<th>Preference Assessment</th>
<th>Treatments Compared</th>
<th>Preference Study Type</th>
<th>Process Variables</th>
<th>Symptom Change</th>
<th>Outcome Variables</th>
<th>Dropout</th>
</tr>
</thead>
<tbody>
<tr>
<td>Elkin et al., 1999</td>
<td>post-hoc from survey</td>
<td>ADM vs. (IPT or CBT)</td>
<td>RCT</td>
<td>alliance (VTAS, BRLI)</td>
<td>-</td>
<td>-</td>
<td>dropout in first 4 weeks</td>
</tr>
<tr>
<td>Rokke et al., 1999</td>
<td>simple preference</td>
<td>SMT target: thoughts / behaviors</td>
<td>choice / no choice</td>
<td>BDI, HRSD, GDI @ post</td>
<td>% remitted @ post</td>
<td>dropout &amp; attendance</td>
<td></td>
</tr>
<tr>
<td>Ward et al., 2000</td>
<td>simple preference</td>
<td>CBT vs. non-directive</td>
<td>partial PRPT / RCT</td>
<td>BDI @ post</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>Lin et al., 2005</td>
<td>simple preference</td>
<td>ADM vs. counseling</td>
<td>assigned treatment</td>
<td>SCL @ 3 &amp; 9 months</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>Chilvers et al., 2007</td>
<td>simple preference</td>
<td>ADM vs. counseling</td>
<td>PRPT</td>
<td>BDI @ post</td>
<td>time to remit; diagnosis @ post</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Leykin et al., 2007</td>
<td>simple preference</td>
<td>ADM vs. CBT</td>
<td>RCT</td>
<td>BDI, HRSD</td>
<td>-</td>
<td>dropout</td>
<td></td>
</tr>
<tr>
<td>Van et al., 2009</td>
<td>refusal to randomize</td>
<td>ADM vs. SPSP</td>
<td>PRPT</td>
<td>remission &amp; response (per HRSD)</td>
<td>-</td>
<td>dropout</td>
<td></td>
</tr>
<tr>
<td>Raue et al., 2009</td>
<td>relative rank + strength of preference</td>
<td>ADM vs. IPT</td>
<td>partial PRPT</td>
<td>HRSD @ post and 24-week follow-up</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>Kocsis et al., 2009</td>
<td>simple preference</td>
<td>ADM vs. CBASP vs. combination</td>
<td>RCT</td>
<td>HRSD</td>
<td>remission &amp; response (per HRSD)</td>
<td>failure to attend; dropout</td>
<td></td>
</tr>
<tr>
<td>Kwan et al., 2010*</td>
<td>simple preference</td>
<td>ADM vs. (BA or CBT)</td>
<td>RCT</td>
<td>HRSD</td>
<td>-</td>
<td>dropout; refuse randomization; attendance</td>
<td></td>
</tr>
</tbody>
</table>

Table 1. Studies of Preference/Outcome Associations in Treatment of Depression.

Note: ADM = anti-depressant medication regimen, *not included in the Swift, Callahan & Vollmer (2011) meta-analysis
Table 2. Summary of Preference/Outcome Findings in Treatment of Depression.

Note: ADM = anti-depressant medication regimen, *not included in the Swift, Callahan & Vollmer (2011) meta-analysis
<table>
<thead>
<tr>
<th></th>
<th>Student Sample</th>
<th></th>
<th>Patient Sample</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(n=198)</td>
<td></td>
<td>(n=38)</td>
<td></td>
</tr>
<tr>
<td>Behavioral</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean (SD)</td>
<td>23.9(4.3)</td>
<td>16.5(4.2)</td>
<td>24.5(4.0)</td>
<td>16.4(4.3)</td>
</tr>
<tr>
<td>Range</td>
<td>11 – 35</td>
<td>5 – 25</td>
<td>17 -35</td>
<td>8 - 25</td>
</tr>
<tr>
<td>Cognitive</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean (SD)</td>
<td>23.5(4.6)</td>
<td>15.9(4.2)</td>
<td>25.8(4.2)</td>
<td>18.6(3.8)</td>
</tr>
<tr>
<td>Range</td>
<td>8 – 35</td>
<td>5 – 25</td>
<td>16 – 34</td>
<td>11 - 25</td>
</tr>
</tbody>
</table>

*Table 3.* Means of Intervention Credibility Ratings and Personal Reactions Ratings of Vignettes & Rationales by Intervention Type and Sample.

Note: ICS = Intervention Credibility Scale, PRR = Personal Reactions to the Rationales
<table>
<thead>
<tr>
<th>Behavioral PRR</th>
<th>Cognitive PRR</th>
<th>Behavioral ICS</th>
<th>Cognitive ICS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Behavioral PRR</td>
<td>-.07</td>
<td>.80*</td>
<td>-.04</td>
</tr>
<tr>
<td>Cognitive PRR</td>
<td>.44*</td>
<td>.11</td>
<td>.72*</td>
</tr>
<tr>
<td>Behavioral ICS</td>
<td>.76*</td>
<td>.40*</td>
<td>.16</td>
</tr>
<tr>
<td>Cognitive ICS</td>
<td>.35*</td>
<td>.81*</td>
<td>.39*</td>
</tr>
</tbody>
</table>

*Table 4. Correlations between PRR and ICS scores, by Sample.*

Note: Correlations for the patient sample (n=38) are listed on the upper right half of the diagonal, and correlations for the student sample (n=193) are on the lower left. Viewing order is partialled out of these correlations. *p < .05
<table>
<thead>
<tr>
<th></th>
<th>Student Sample (n = 198)</th>
<th>Patient Sample (n = 38)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>ICS-BvC (difference score)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean (SD)</td>
<td>.44 (4.4)</td>
<td>-1.32 (5.5)</td>
</tr>
<tr>
<td>Range</td>
<td>-14 to 14</td>
<td>-16 to 8</td>
</tr>
<tr>
<td><strong>PRR-BvC (difference score)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean (SD)</td>
<td>.01 (4.5)</td>
<td>-2.70 (5.3)</td>
</tr>
<tr>
<td>Range</td>
<td>-10 to 12</td>
<td>-14 to 6</td>
</tr>
</tbody>
</table>

*Table 5. Means of ICS and PRR Difference Scores, by Sample.*

Note: Scores are scaled so that higher values represent higher relative scores for behavioral interventions as compared to cognitive interventions.
<table>
<thead>
<tr>
<th></th>
<th>Student Sample</th>
<th></th>
<th>Patient Sample</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(n = 198)</td>
<td></td>
<td>(n=38)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>t</td>
<td>p</td>
<td>r</td>
<td>t</td>
</tr>
<tr>
<td>ICS Behavioral</td>
<td>-.13</td>
<td>.89</td>
<td>-.00</td>
<td></td>
</tr>
<tr>
<td>ICS Cognitive</td>
<td>-.53</td>
<td>.60</td>
<td>-.04</td>
<td></td>
</tr>
<tr>
<td>ICS-BvC</td>
<td>-.64</td>
<td>.52</td>
<td>-.05</td>
<td></td>
</tr>
<tr>
<td>PRR Behavioral</td>
<td>-2.15</td>
<td>.03</td>
<td>-.15</td>
<td></td>
</tr>
<tr>
<td>PRR Cognitive</td>
<td>.69</td>
<td>.49</td>
<td>.05</td>
<td></td>
</tr>
<tr>
<td>PRR-BvC</td>
<td>-1.31</td>
<td>.19</td>
<td>-.09</td>
<td></td>
</tr>
</tbody>
</table>

*Table 6.* RFD Total scores as Predictors of Intervention ICS and PRR Ratings and Difference Scores, with Viewing Order Co-varied, in Student and Patient Samples.

ICS = Intervention Credibility Scale, PRR = Personal Reactions to the Rationales, RFD = Reasons for Depression Inventory, BvC suffix denotes a difference score, where higher scores reflect relatively higher behavioral ratings.
### Table 7.

RFD Total Scores and Intervention ICS and PRR Difference Scores in the prediction of Treatment Preference scores in the Student Sample.

ICS = Intervention Credibility Scale, PRR = Personal Reactions to the Rationales, RFD = Reasons for Depression Inventory, BvC suffix denotes a difference score, where higher scores reflect relatively higher behavioral ratings.
<table>
<thead>
<tr>
<th></th>
<th>Separate Predictors (df = 38)</th>
<th>Both Predictors in Model (df = 37)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>( t ) ( p ) ( r )</td>
<td>( t ) ( p ) ( r )</td>
</tr>
<tr>
<td>RFD Total</td>
<td>-2.56 .02 -.40</td>
<td>-1.29 .21 -.22</td>
</tr>
<tr>
<td>ICS-BvC</td>
<td>7.32 &lt;.01 .78</td>
<td>6.48 &lt;.01 .74</td>
</tr>
<tr>
<td>PRR Model</td>
<td></td>
<td></td>
</tr>
<tr>
<td>RFD Total</td>
<td>-2.56 .02 -.40</td>
<td>-.74 .46 -.13</td>
</tr>
<tr>
<td>PRR-BvC</td>
<td>7.82 &lt;.01 .80</td>
<td>6.79 &lt;.01 .76</td>
</tr>
</tbody>
</table>

Table 8. RFD Total Scores and Intervention ICS and PRR Difference Scores in the prediction of Treatment Preference scores in the Patient Sample.

ICS = Intervention Credibility Scale, PRR = Personal Reactions to the Rationales, RFD = Reasons for Depression Inventory, BvC suffix denotes a difference score, where higher scores reflect relatively higher behavioral ratings.
<table>
<thead>
<tr>
<th></th>
<th>Overall</th>
<th>Behavioral</th>
<th>Cognitive</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Intake</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>n</td>
<td>38</td>
<td>19</td>
<td>19</td>
</tr>
<tr>
<td>HRSD</td>
<td>20.2 (5.2)</td>
<td>20.3 (6.2)</td>
<td>20.2 (4.2)</td>
</tr>
<tr>
<td><strong>Week 4</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>n</td>
<td>32</td>
<td>15</td>
<td>17</td>
</tr>
<tr>
<td>HRSD</td>
<td>17.9 (6.6)</td>
<td>17.1 (7.8)</td>
<td>18.6 (5.6)</td>
</tr>
<tr>
<td>DSI</td>
<td>28.9 (2.9)</td>
<td>28.9 (3.0)</td>
<td>28.9 (2.8)</td>
</tr>
<tr>
<td><strong>Week 8</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>n</td>
<td>29</td>
<td>15</td>
<td>14</td>
</tr>
<tr>
<td>HRSD</td>
<td>14.3 (6.7)</td>
<td>14.5 (7.9)</td>
<td>14.1 (5.3)</td>
</tr>
<tr>
<td>DSI</td>
<td>59.3 (4.3)</td>
<td>61.3 (4.6)</td>
<td>57.3 (2.8)</td>
</tr>
</tbody>
</table>

Table 9. HRSD Data Availability, Means, and Average Time to Assessment, Overall and by Treatment Condition.

DSI = days since intake, HRSD = Hamilton Rating Scale for Depression. Note: The target value of days since intake for the week 4 assessment is 28 days, and for the week 8 assessment is 56 days.
<table>
<thead>
<tr>
<th></th>
<th>Overall (n = 38)</th>
<th>Received Behavioral (n=19)</th>
<th>Received Cognitive (n=19)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Prefer Behavioral</td>
<td>10 (26%)</td>
<td>4 (21%)</td>
<td>6 (32%)</td>
</tr>
<tr>
<td>Prefer Cognitive</td>
<td>19 (50%)</td>
<td>11 (58%)</td>
<td>8 (42%)</td>
</tr>
<tr>
<td>Neutral Preference</td>
<td>9 (24%)</td>
<td>4 (21%)</td>
<td>5 (26%)</td>
</tr>
<tr>
<td>Matched</td>
<td>21 (55%)</td>
<td>8 (43%)</td>
<td>13 (68%)</td>
</tr>
<tr>
<td>Mismatched</td>
<td>17 (45%)</td>
<td>11 (57%)</td>
<td>6 (32%)</td>
</tr>
</tbody>
</table>

*Table 10. Categorical Preference Ratings and Treatment Matching Status.*
### Table 11. Demographic and Treatment Characteristics by Treatment Condition, with Group Comparisons.

<table>
<thead>
<tr>
<th></th>
<th>Behavioral (n=19)</th>
<th>Cognitive (n=19)</th>
<th>Group Comparison</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>34.4(15.8)</td>
<td>31.9(14.9)</td>
<td>( t(36) = .51, \ p = .61 )</td>
</tr>
<tr>
<td>Female</td>
<td>8(42%)</td>
<td>13 (68%)</td>
<td>( \chi^2 = 2.66, \ p = .10 )</td>
</tr>
<tr>
<td>Minority</td>
<td>2 (11%)</td>
<td>5 (26%)</td>
<td>Fisher’s exact, ( p = .40 )</td>
</tr>
<tr>
<td>Single</td>
<td>16 (84%)</td>
<td>14 (74%)</td>
<td>Fisher’s exact, ( p = .69 )</td>
</tr>
<tr>
<td>Chronic/dysthmic</td>
<td>9 (47%)</td>
<td>10(53%)</td>
<td>( \chi^2 = .11, \ p = .75 )</td>
</tr>
<tr>
<td>Prior episodes</td>
<td>2.7(1.7)</td>
<td>2.8(1.8)</td>
<td>( t(32) = -.19, \ p = .85 )</td>
</tr>
<tr>
<td>Comorbid Anxiety Disorder</td>
<td>10 (52%)</td>
<td>13 (68%)</td>
<td>( \chi^2 = .99, \ p = .31 )</td>
</tr>
<tr>
<td>Comorbid Axis II Diagnosis</td>
<td>1 (5%)</td>
<td>3 (16%)</td>
<td>( \chi^2(1) = 1.1, \ p = .29 )</td>
</tr>
<tr>
<td>Number of Sessions</td>
<td>6.2 (3.0)</td>
<td>6.8 (2.1)</td>
<td>( t(36) = -.76, \ p = .45 )</td>
</tr>
<tr>
<td>On concurrent medication</td>
<td>5 (26%)</td>
<td>4 (21%)</td>
<td>Fisher’s exact, ( p = .71 )</td>
</tr>
</tbody>
</table>

Note: Prior episode data was available for 17 subjects in each group. Tests reported in group comparison include \( t \)-tests for continuous variables and Chi square tests for categorical variables and Fisher’s exact tests where applicable.
<table>
<thead>
<tr>
<th></th>
<th>Interaction with Time (df = 28)</th>
<th>Interaction with Time and Intervention (df=28)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>t</td>
<td>p</td>
</tr>
<tr>
<td>RFD total score</td>
<td>1.23</td>
<td>.23</td>
</tr>
<tr>
<td>ICS-BvC</td>
<td>.09</td>
<td>.93</td>
</tr>
<tr>
<td>PRR-BvC</td>
<td>.46</td>
<td>.65</td>
</tr>
<tr>
<td>Preference rating</td>
<td>.34</td>
<td>.74</td>
</tr>
<tr>
<td>Treatment match</td>
<td>-.36</td>
<td>.72</td>
</tr>
</tbody>
</table>

*Table 12. Potential Predictors of Slope of HRSD Symptom Change across the First Eight Weeks of Treatment After Covarying Medication Status.*

ICS = Intervention Credibility Scale, PRR = Personal Reactions to the Rationales, RFD = Reasons for Depression Inventory, BvC suffix denotes a difference score, where higher scores reflect relatively higher behavioral ratings.
<table>
<thead>
<tr>
<th></th>
<th>Dropouts (n=9)</th>
<th>Completer (n=29)</th>
<th>Group Comparison</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>33.4 (15.2)</td>
<td>33.1 (15.4)</td>
<td>$t(36) = -0.06, p = .94$</td>
</tr>
<tr>
<td>Female</td>
<td>4 (44%)</td>
<td>17 (59%)</td>
<td>Fisher’s exact, $p = .70$</td>
</tr>
<tr>
<td>Minority</td>
<td>3 (33%)</td>
<td>4 (14%)</td>
<td>Fisher’s exact, $p = .32$</td>
</tr>
<tr>
<td>Single</td>
<td>7 (78%)</td>
<td>23 (79%)</td>
<td>Fisher’s exact, $p = .99$</td>
</tr>
<tr>
<td>Chronic/dysthymic</td>
<td>7 (37%)</td>
<td>2 (11%)</td>
<td>Fisher’s exact, $p = .12$</td>
</tr>
<tr>
<td>Prior episodes</td>
<td>2.6 (1.8)</td>
<td>2.8 (1.8)</td>
<td>$t(32) = .41, p = .69$</td>
</tr>
<tr>
<td>Comorbid Anxiety Disorder</td>
<td>5 (56%)</td>
<td>18 (62%)</td>
<td>Fisher’s exact, $p = .99$</td>
</tr>
<tr>
<td>Comorbid Axis II Diagnosis</td>
<td>1 (11%)</td>
<td>3 (10%)</td>
<td>Fisher’s exact, $p = .99$</td>
</tr>
<tr>
<td>Number of Sessions</td>
<td>3.3 (2.4)</td>
<td>7.5 (1.6)</td>
<td>$t(36) = 6.02, p &lt;.0001$</td>
</tr>
<tr>
<td>On concurrent medication</td>
<td>3 (33%)</td>
<td>6 (21%)</td>
<td>Fisher’s exact, $p = .66$</td>
</tr>
</tbody>
</table>

*Table 13. Demographic and Treatment Characteristics by Dropout Status.*

Prior episode data was available for 25 subjects in the non-dropout group, and all dropouts. Tests reported under Group Comparison include t-tests for continuous variables and Fisher’s exact tests for categorical variables due to the small number of dropouts making Chi square analysis inappropriate.
<table>
<thead>
<tr>
<th>Predictor</th>
<th>Single Predictor</th>
<th>Interaction with Intervention</th>
<th>Interaction with Intervention</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\chi^2$</td>
<td>$p$</td>
<td>OR (95% CI)</td>
</tr>
<tr>
<td>RFD total score</td>
<td>3.0</td>
<td>.08</td>
<td>2.6 (.88-7.7)</td>
</tr>
<tr>
<td>ICS-BvC</td>
<td>1.6</td>
<td>.20</td>
<td>.60 (.27 – 1.3)</td>
</tr>
<tr>
<td>PRR-BvC</td>
<td>.18</td>
<td>.67</td>
<td>.84 (.37 – 1.9)</td>
</tr>
<tr>
<td>Preference rating</td>
<td>4.8</td>
<td>.03</td>
<td>.49 (.26 - .93)</td>
</tr>
<tr>
<td>Treatment match</td>
<td>.40</td>
<td>.52</td>
<td>1.7 (.34 – 8.5)</td>
</tr>
</tbody>
</table>

Table 14. Potential Predictors of Dropout Across the First Eight Weeks of Treatment.

Note: All analyses include intake HRSD and concurrent medication status as covariates. Odds ratios reported by condition are taken from analyses restricted to that condition. ICS = Intervention Credibility Scale, PRR = Personal Reactions to the Rationales, RFD = Reasons for Depression Inventory, BvC suffix denotes a difference score, where higher scores reflect relatively higher behavioral ratings.
Appendix B: Figures
Figure 1. RFD Score by Treatment Assignment in the Prediction of Dropout.

Note: RFD scores are standardized residuals created to account for effects of viewing order. Higher PRR-BvC scores reflect relatively higher ratings of behavioral treatment. Intake HRSD and medication status were covariates in this model.
Figure 2. PRR-BvC Score by Treatment Assignment in the Prediction of Dropout

Note: PRR–BvC scores are standardized residuals created to account for effects of viewing order. Higher PRR-BvC scores reflect relatively higher ratings of behavioral treatment. Intake HRSD and medication status were covariates in this model.