The Role of Interpersonal Vulnerability in Cognitive Therapy for 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

Katherine Emma Sasso, M.A.
Graduate Program in Psychology

The Ohio State University
2017

Dissertation Committee:
Professor Daniel Strunk, Advisor
Professor Jennifer Cheavens
Professor Amelia Aldao
Abstract

Interpersonal problems and related maladaptive personality traits have been found to predict significant variability in symptom change, dropout, and processes of change in treatments for depression and more specifically Cognitive Therapy (CT; Diener & Monroe, 2011; Hardy et al., 2001; Newton-Howes et al., 2006; McEvoy, Burgess, & Nathan, 2013; Renner et al., 2012; Swift & Greenberg, 2012); however, few studies have examined clients’ interpersonal vulnerabilities as perceived by others, with the majority of studies to date relying exclusively on self-report or diagnostic information. How one is perceived by others may be more consequential than one’s self-views in a treatment setting, as others’ ratings of target individuals’ personality and interpersonal tendencies have demonstrated stronger predictive validity than self-report ratings in a variety of settings, including some in the clinical realm (Connelly & Ones, 2010; Klein, 2003). While little work has examined the predictive validity of observer rated interpersonal vulnerability in a treatment context, how clients’ are perceived by others may explain meaningful variability in their therapeutic experience.

The Thin Slice-observer (TS) and Standardized Interaction Task (SIT) measures used in previous studies provide ecologically valid assessments of clients’ interpersonal vulnerabilities as perceived by others. The TS measure, which gauges others’ perceptions of clients’ personality traits and interpersonal tendencies from brief (<5 min) video-clips
of clients talking about themselves, has shown predictive validity above and beyond commonly-used self-report measures in predicting symptom change and dropout in CT for depression specifically (Sasso & Strunk, 2013). The SIT measure, which gauges others’ perceptions of a target individual’s performance on video-recorded role-play tasks, has been found to measure variability in interpersonal functioning not captured by related self-report information (Leising et al. 2011).

I aimed to compliment traditional self-report measures with both SIT and TS ratings in order to provide a more comprehensive view of how clients’ interpersonal vulnerabilities and related personality traits predict CT processes and outcomes in a sample of 126 clients participating in a course of CT for depression. Overall, findings supported the notion that others’ perceptions of interpersonal deficits and related personality traits have predictive validity the clinical realm. Specifically, therapist thin slice ratings of heightened client neuroticism, SIT-observer ratings indicative of a more favorable role play performance, and heightened self-reported personality pathology significantly predicted a steeper slope of subsequent symptom reduction. While I did not find that any therapeutic processes mediated these relations, I did find significant main effects of interpersonal predictors on therapeutic processes. Thin slice observer ratings of clients’ pro-treatment behaviors (i.e., likeability, agreeableness, conscientiousness and, communication skills) predicted formation of the therapeutic alliance and client reported in-session cognitive changes, with clients perceived to be lower in pro-treatment behaviors experiencing less favorable levels of these therapeutic processes. Therapists’ perceptions of client sociability also showed a significant positive relation to the
subsequently formed therapeutic alliance. Lastly, clients who reported heightened personality pathology tended to experience more in-session cognitive change early in treatment. These findings suggest that a multi-modal approach to assessing interpersonal vulnerability may help clinicians to better identify at-risk clients.
Acknowledgements

I would first like to acknowledge the contributions of my advisor, Dr. Dan Strunk, without whom this project would not have been possible. I am grateful for the guidance and support Dr. Strunk has offered me on this project and all of my research endeavors over the past nine years in our work together. I would also like to thank my committee members, Dr. Jen Cheavens and Dr. Amelia Aldao for their contributions, time, and support throughout this process.

I would also like to acknowledge my colleagues in the Depression Research Lab, particularly the therapists and assessors who treated clients in the study under investigation. Matt, Erin, Ben, and Justin; your countless late hours and hard work were all integral to the success of this project. I would specifically like to thank Justin Braun, who coordinated several aspects of this treatment study and has collaborated with me and supported me through several of my research projects over the past nine years.

Finally, I want to express my love and gratitude for my family and friends. I am so thankful for my parents and brothers, who have always provided me with unconditional love and encouragement while modeling a strong work ethic. I also wish to thank my significant other, Ben, and close friends for their support and patience throughout this process.
Vita

2008.................................................................B.S., Psychology (Summa Cum Laude), The Ohio State University

2012-2013 ......................................................University Fellow, The Ohio State University

2013-2014 ........................................................Graduate Teaching Associate, Department of Psychology, The Ohio State University

2014...............................................................M.A., Psychology, The Ohio State University

2015...............................................................Doctoral Candidate, Department of Psychology, The Ohio State University

2015-present...................................................

Graduate Research Associate, Nationwide Center for Advanced Customer Insights, The Ohio State University
Publications


Fields of Study

Major Field: Psychology
# Table of Contents

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

Acknowledgments........................................................................................................... v  

Vita................................................................................................................................... vi  

List of Tables ................................................................................................................... xiii  

List of Figures............................................................................................................... xv  

Chapter 1: Introduction .................................................................................................... 1  
  Overview of CT ........................................................................................................ 2  
  Prior Work on the Role of Interpersonal Vulnerability in Depression and CT ........ 5  
  Interpersonal vulnerability in depression ............................................................... 5  
  Interpersonal vulnerability and associated traits as predictors of depression treatment outcome .............................................................. 7  

Key Processes in CT: Relations to Interpersonal Vulnerability and Outcome .......... 9  
  Main effects of CT processes on outcome ............................................................... 9  
  The alliance and interpersonal vulnerability ........................................................ 14  
  Cognitive change and interpersonal vulnerability .............................................. 15  
  Cognitive change and the alliance as mediators: Preliminary evidence .......... 16  

Ways of Measuring Interpersonal Vulnerability and Associated Traits ............ 19  
  Common self-report and assessor-administered measures ............................... 19  

viii
Self-report interpersonal vulnerability: IIP-32 and its predecessors ..........22
Over-Reliance on Self-Reported Interpersonal Vulnerability ..............25
Means of Assessing Others’ Perceptions of Interpersonal Vulnerability ........26
Performance on interpersonal role plays .............................................29
Purpose of This Study ........................................................................31
Hypotheses ..........................................................................................32
Objective 1: Hypotheses 1A-1B and 2A-2B ........................................32
Objective 2: Hypotheses 3-4 .................................................................33
Objective 3 ..........................................................................................34
Chapter 2: Methods ..............................................................................35
Data Collection and Cleaning ..............................................................35
Participants ...........................................................................................35
Demographics and descriptive statistics ..........................................37
Diagnoses and symptoms at intake ..................................................37
Assessors and Therapists .................................................................38
Session Frequency ..............................................................................38
Measures .............................................................................................39
Diagnoses ............................................................................................39
Outcome measures ............................................................................39
Dropout ...............................................................................................39
Depressive symptoms ........................................................................39
Therapeutic process measures .......................................................40
Objective 1: Interpersonal Vulnerability Predicting Treatment Outcome .......... 58
  Hypothesis 1A ................................................................. 58
  Hypothesis 1B ................................................................. 59
  Hypothesis 2A ................................................................. 59
  Hypothesis 2B ................................................................. 60
  Exploratory analyses: Informant differences ........................................ 61
Objective 2: Interpersonal Vulnerability Predicting Key CT Processes .......... 62
  Hypothesis 3 ..................................................................... 62
  Hypothesis 4 ..................................................................... 63
  Exploratory analyses ................................................................... 65
    Predictive effects of therapist ratings on average alliance ................. 65
    Predictive effects of therapist ratings on clients’ average in-session cognitive change ................................................................. 66
    Variability in effects across TS informants .................................. 67
Objective 3: Alliance and In-session Cognitive Change as Mediators ............ 67
Chapter 4: Discussion ........................................................................... 72
  Key Predictors: Psychometric Properties and Relations among Constructs .... 73
    Distinct constructs discernable to thin slice raters ......................... 73
    Construct validity and relations among predictors ............................. 74
Main Findings and Support for Hypotheses ........................................ 76
  Objective 1: Interpersonal vulnerability predicting slope of symptom change ... 76
List of Tables

Table 1. Primary and Comorbid Diagnoses and Specifiers at Intake .......................106
Table 2. Rotated Factor Pattern from Exploratory Factor Analysis for Thin Slice-Observer Measure ............................................................................................................107
Table 3. Correlations Among Candidate Predictors and Self-Reported Baseline Depressive Symptoms ..........................................................................................................................108
Table 4. Means on Key Predictors as a Function of Comorbid Anxiety Diagnoses ......109
Table 5. Parameter Estimates from Mixed Model Examining Effects of Interpersonal Vulnerabilities on Slope of Symptom Change .................................................................110
Table 6. Observer-rated and Self-Reported Interpersonal Vulnerability Predicting Dropout (Controlling for Intake Depressive Symptoms) .................................................111
Table 7. Parameter Estimates from Mixed Model Examining Effects of Therapist Thin Slice Ratings on Slope of Symptom Change ........................................................................112
Table 8. Therapist Thin Slice Ratings Predicting Dropout (Controlling for Intake Depressive Symptoms) ..................................................................................................................113
Table 9. Parameter Estimates from Mixed Model Examining Effects of Interactions between Thin Slice Constructs and Rater on Subsequent Symptom Change ...............114
Table 10. Parameter Estimates from Mixed Model Examining Effects of Observer-Rated Interpersonal Vulnerability on Average Alliance Scores and Slope of Alliance ........115
Table 11. Parameter Estimates from Mixed Model Examining Effects of Thin Slice Observer: Pro-treatment Behaviors on Alliance while Controlling for Potential Confounds ........................................................................................................................116

Table 12. Parameter Estimates from Mixed Model Examining Effects of Observer-Rated Interpersonal Vulnerability on Average Cognitive Change and Slope of Cognitive Change .............................................................................................................................117

Table 13. Parameter Estimates from Mixed Model Examining Effects of Thin Slice Observer: Pro-treatment Behaviors on Cognitive Change while Controlling for Potential Confounds ..........................................................................................................................118

Table 14. Parameter Estimates from Mixed Model Examining Effects of Therapist-Rated Interpersonal Vulnerability on Average Alliance Scores and Slope of Alliance ..........................................................119

Table 15. Parameter Estimates from Mixed Model Examining Effects of Thin Slice Therapist: Sociability on Alliance while Controlling for Potential Confounds .........................120

Table 16. Parameter Estimates from Mixed Model Examining Effects of Therapist-Rated Interpersonal Vulnerability on Average In-session Cognitive Change Scores and Slope of In-session Cognitive Change .........................................................................................................................121

Table 17. Parameter Estimates from Mixed Model Examining Effects of Interactions between Thin Slice Constructs and Rater on Alliance .............................................................................................................................122

Table 18. Parameter Estimates from Mixed Model Examining Effects of Interactions between Thin Slice Constructs and Rater on In-session Cognitive Change .................................123
List of Figures

Figure 1. Scree Plot of Eigenvalues Derived from Parallel Analysis of Thin Slice Observer Measure .............................................................................................................125

Figure 2. Scree Plot of Eigenvalues Derived from Parallel Analysis of Thin Slice Therapist Measure .............................................................................................................126

Figure 3. Scree Plot of Eigenvalues Derived from Parallel Analysis of In-Session Cognitive Change Measure .............................................................................................................127
Chapter 1: Introduction

Cognitive Therapy (CT) has been established as an efficacious treatment for depression, with several studies supporting its efficacy in reducing depressive symptoms during acute treatment and maintaining symptom improvement post-treatment (for a review, see Hollon, Stewart, & Strunk, 2006; Strunk & DeRubeis, 2001). However, clients show considerable variability in their response to treatment, with recovery and remission occurring typically in 40-60% of CT treated clients with major depressive disorder (MDD; DeRubeis et al., 2005; Hollon, Thase, & Markowitz, 2002). Thus, not all clients achieve optimal outcomes and there is still an unmet need among a sizable portion of clients receiving CT. CT clients with increased interpersonal problems have been found to experience less depressive symptom improvement and to be at a higher risk of attrition than those with less marked interpersonal vulnerabilities (Hardy et al., 2001; McEvoy, Burgess, & Nathan, 2013). To the extent that interpersonal vulnerability is related to key CT processes, such as cognitive change and the therapeutic alliance, it is possible that such processes may partially account for the relation between interpersonal problems and suboptimal outcomes (Holtforth et al., 2014; Levin, Henderson, & Ehrenreich-May, 2012; McEvoy, Burgess, & Nathan, 2013).

Interpersonal vulnerabilities are assessed primarily via client self-report, leaving it unclear whether others’ perceptions of such vulnerabilities may also predict meaningful
variations in treatment response or therapeutic mechanisms. Others’ perceptions of clients’ interpersonal deficits may be important, to the extent that they are related to the difficulties clients encounter when interacting with others (e.g., friends, therapists, partners) and implementing skills learned in treatment in their day-to-day life. Others’ ratings of clients maladaptive personality traits have been shown to predict subsequent symptom change and dropout in a sample of clients who participated in CT for depression, above and beyond traditional self-report measures (Sasso & Strunk, 2013). This study aims to clarify the relation between interpersonal vulnerabilities, treatment outcomes, and key processes in a sample of clients participating in CT for depression, and expands upon prior work by assessing clients’ interpersonal vulnerability as perceived by others. The following sections provide an overview of the relation between interpersonal vulnerability and depression, the effects of interpersonal vulnerability on treatment outcome and treatment mechanisms, and the benefits of assessing such vulnerability from others’ perspectives. To start, a basic overview of the primary treatment of focus in this study, CT for depression, is provided.

**Overview of CT**

Cognitive therapy is defined as a structured, collaborative, and time-limited approach to treating depression (Beck, Rush, Shaw, & Emery, 1979). CT is based primarily on the theoretical notion that depressed individuals hold negative beliefs regarding themselves, the world, and the future, which in turn have a detrimental impact on their affect and behavior. Thus, clients in CT learn both cognitive and behavioral skills that allow them to respond more effectively to heightened negative affect. Cognitive
techniques encourage clients to identify and systematically challenge and correct maladaptive and distorted negative automatic thoughts. Therapists also work with clients to highlight and target biases in information processing underlying clients’ everyday experiences and behaviors (Beck et al., 1979). Behavioral experiments, which often require clients to interact with others, may be conducted to test clients’ negative beliefs. CT also includes some techniques targeted more specifically at behavioral change. Therapists work with clients to increase client engagement in behaviors likely to produce a sense of pleasure and mastery, target patterns of avoidance that contribute to depressive symptoms, and build problem solving skills. Clients are taught to recognize the relationship between their mood and activities and to re-structure their time accordingly in order to address periods of low mood. They are also taught how to break challenging situations down into smaller pieces in order to overcome obstacles.

In the first session of CT, clients are typically introduced to the treatment rationale and therapeutic techniques. One such technique is the use of homework assignments (e.g. activity logs, thought records, behavioral experiments), which function to reinforce therapeutic content by aiding clients in the practice and independent use of therapeutic skills in the time between sessions. Additional early sessions focus on getting clients back to their premorbid level of functioning by incorporating assignments designed to re-engage clients in activities they previously found pleasurable. Simultaneously, therapists introduce cognitive techniques, with a primary one being thought records, which aid clients in the observation and recording of their automatic thoughts, or spontaneous cognitive assessments of an event. Thought records are then
reviewed in session to reality-test and correct negative automatic thoughts. In later sessions, client and therapist discuss cognitions more in depth, in an attempt to identify recurrent or common themes (i.e., schemas). Clients also begin working independently on recognizing cognitive errors and reviewing alternative explanations for negative automatic thoughts. In the final sessions, the end goal is for clients to achieve autonomy by effectively serving as their own therapist.

Many clients in CT experience significant reductions in depressive symptoms, with an intent to treat analysis examining client outcomes from four studies of CT for depression finding a large average pre- to post-treatment Cohen’s d effect size of approximately 3.0 (see review by DeRubeis, Gelfand, Tang, & Simons, 1999) on the interviewer-administered Hamilton Rating Scale for Depression (HRSD; Hamilton, 1960). Despite encouraging effect sizes for CT on average, clients with increased interpersonal vulnerability tend to show less robust response to treatment, as discussed below. Therapeutic process variables, or key mechanisms of change in CT, may, in part, account for such relations to the extent that such processes are negatively influenced by clients’ heightened interpersonal vulnerability. This study aims to elucidate the role of interpersonal vulnerability in CT for depression by exploring the utility of others’ perceptions of clients’ interpersonal traits in predicting key CT processes and outcome. Current research examining the role of interpersonal vulnerability in depression and CT is reviewed in the following sections. In later sections findings on key mechanisms of change in CT and potential relations between these and interpersonal vulnerabilities of interest are discussed.
Prior Work on the Role of Interpersonal Vulnerability in Depression and CT

Interpersonal vulnerability in depression. MDD consists of a variety of somatic, cognitive, affective, and behavioral symptoms that impact not only the way the individual feels and thinks about themselves but also the way in which they interact with the people in their environment (Hames, Hagan, & Joiner, 2013; Holtforth et al., 2014).

The relations between interpersonal vulnerabilities and depression have been well-studied, with several maladaptive interpersonal processes being identified as risk factors and correlates of depression. Broadly, chronic interpersonal stress and major interpersonal stressful life events have been identified as statistically unique predictors of risk for onset of a major depressive episode (MDE), over and above other correlated forms of non-interpersonal life stress (Vrshek-Schallhorn et al., 2015). Depressive symptoms have also been found to predict pessimistic biases in predictions of how one will be perceived by others (Strunk & Adler, 2009).

Beyond these broad relations between depressive symptoms and stressful life events, more specific interpersonal behavioral tendencies and associated personality traits have been related to depression. For example, lack of assertiveness has been found to be a significant predictor of depression (Hames et al., 2013; Holtforth et al., 2014). Likewise, those with depression have a tendency to subjugate themselves by prioritizing others’ needs over their own (McEvoy, Burgess, Page, Nathan, & Fursland, 2013).

Maladaptive attachment styles, or an infants’ attachment to primary caregivers thought to influence future relationships, have also been related to depression. Specifically, insecure attachment styles (i.e., anxious, avoidant, and fearful) have been found to be more
common among those with depressive diagnoses (for a review see, Hames et al., 2013), and insecurely attached adults tend to have more difficulty in forming and maintaining close personal relationships (Horowitz, Rosenberg, & Bartholomew, 1993). Individuals with depression have also been found to engage in both excessive reassurance and negative feedback seeking behaviors. Excessive reassurance seeking refers to the tendency to persistently seek assurance from others that one is lovable and worthy, regardless of whether such assurance has already been provided; whereas, negative feedback seeking refers to the tendency to solicit criticism from others in an attempt at self-verification (Hames et al., 2013). Both excessive reassurance and negative feedback seeking have been found repeatedly to be perceived negatively by others and to elicit social rejection (Joiner, Katz, & Lew, 1997; Joiner, Alfano, & Metalsky, 1992; Starr & Davila, 2008)

Even quickly detectable basic behavioral differences, which may contribute to interpersonal difficulties, have been identified between depressed and non-depressed individuals. The voices of individuals with depression have been found to be perceived more negatively than that of their non-depressed counterparts, with depressed individuals tending to speak more slowly, with lower voice modulation and at lower volumes than those without depression (Hames et al., 2013). Other quickly detectable behavioral traits associated with depression include reduced eye contact, exaggerated facial expressions of sadness, downward gaze, and limited gestures (Hames et al., 2013). Taken together, poor social skills, excessive reassurance seeking, maladaptive attachment styles, and negative verbal and nonverbal behaviors contribute to the heightened interpersonal difficulties
associated with depression. The following sections review how such difficulties have been found to impact treatment outcomes, including symptom change and early attrition, or dropout. While findings from studies examining different treatment modalities are included, reporting has been limited mostly to those studies that examine depression as a primary target.

**Interpersonal vulnerability and associated traits as predictors of depression treatment outcome.** Despite the relation between interpersonal vulnerability and depressive symptoms, associations among the two are not frequently reported in the CT for depression literature (McEvoy, Burgess, Page, Nathan, & Fursland, 2013) and studies that do examine these associations rely almost exclusively on self-reported interpersonal vulnerability as assessed using one of the 10 variants of the Inventory of Interpersonal Problems (IIP; for details see below and Horowitz, Rosenberg, Baer, Ureño, & Villaseñor, 1988). A few studies examining predictors of outcome in CT for depression have found that self-reported interpersonal problems predict post-treatment depressive symptoms, with greater pre-treatment interpersonal problems predicting significantly higher post-treatment depressive symptoms after controlling for pre-treatment symptom severity (Hardy et al., 2001; McEvoy, Burgess, & Nathan, 2013). One study found that increased pre-treatment interpersonal distress predicted less marked reductions in depressive symptoms over the course of treatment (Renner et al., 2012). Heightened interpersonal vulnerability has also been found to predict increased risk of early attrition from CT for depression (McEvoy, Burgess, & Nathan, 2013; McEvoy et al., 2014). Specifically, McEvoy and colleagues (2013) found that clients who reported ‘finding it
hard to be supportive of others’ and ‘not being open about problems’ had significantly higher attrition rates.

Similar to interpersonal distress, the presence of personality disorders (PDs) or maladaptive variants of personality traits have been shown to predict suboptimal outcomes across a variety of treatments for depression (Gorwood et al., 2010; Joyce et al., 2007; Levenson, Wallace, Fournier, Rucci, & Frank, 2012; Wardenaar, Conradi, Bos, & de Jonge, 2014), with a recent meta-analysis reporting the risk of poor outcome (i.e., no recovery) significantly increased among those with a PD (~55%) compared to those without (~45%; Newton-Howes, Tyrer, & Johnson, 2006). Indeed interpersonal impairment and distress is thought to be a core element of most PDs, and suboptimal outcomes have also been demonstrated among those with PDs and PD like traits in CT for depression specifically (Bagby et al., 2008; Fournier et al., 2008; Sasso & Strunk, 2013; Spek, Nyklíček, Cuijpers, & Pop, 2008). PDs have also been associated with higher risk of treatment dropout in both in-patient and out-patient treatment settings (see Swift & Greenberg, 2012). Taken together, both interpersonal distress and PD traits have been associated with suboptimal treatment outcomes across a variety of treatment modalities and within CT for depression specifically.

The above findings suggest that interpersonal vulnerability and associated personality traits at baseline predict significant variability in depressive symptom change and dropout over the course of therapy; however, it remains less clear exactly how interpersonal vulnerability comes to affect these outcomes. Any given CT process, or mechanism of change, may reflect a possible means by which interpersonal
vulnerabilities at baseline exert their observed effects on outcome in depression treatment. Mediator variables can explain how or why a given predictor affects outcome (Baron & Kenny, 1986). According to the criteria laid out by Kraemer and colleagues (2002), three conditions must be met to support a mediational relationship between pre-existing interpersonal vulnerabilities, CT processes, and outcome: (1) All three variables must be significantly related to one another; (2) Interpersonal vulnerability must temporally precede measurement and occurrence of the given CT process and the measured outcome, and; (3) A significant relationship between interpersonal vulnerability and treatment outcome must be accounted for (i.e., attenuated) when the CT process is included in the model (Kraemer, Wilson, Fairburn, & Agras, 2002). The above sections provide support for the relation between interpersonal vulnerability and treatment outcome. Below, evidence supporting the relation between CT processes, outcome, and interpersonal vulnerability is reviewed, as well as preliminary findings supporting the notion that CT processes may account for significant relations between interpersonal vulnerability and treatment outcomes.

**Key Processes in CT: Relations to Interpersonal Vulnerability and Outcome**

**Main effects of CT processes on outcome.** Key CT processes, or proposed mechanisms of change, may potentially account for the relationship between baseline interpersonal vulnerability and treatment outcomes, to the extent that they predict such outcomes and are impacted by interpersonal deficits. Theoretically, one might predict that a client with heightened interpersonal vulnerability has a more difficult time acquiring the skills that such mechanisms are intended to teach, to the extent that his/her interpersonal
deficits impede out of session practice, real-world applications of such skills, and formation of a strong working relationship with his/her therapist. With regards to the empirical findings in support of a mediational theory, specifically criteria one, empirical evidence suggests that two key CT processes, the therapeutic alliance and cognitive change, have shown small to moderate relations to outcome in CT for depression. The therapeutic alliance refers to the affective bond formed between client and therapist and is typically assessed early in treatment or at the midpoint through client, therapist, or observer ratings. One of the most common alliance rating measures used is the short form of the Working Alliance Inventory (WAI, see below for a brief discussion of psychometric properties; Horvath & Greenberg, 1989; Tracey & Kokotovic, 1989).

Cognitive change is a term often used to refer to reductions in the strength and frequency of clients’ negative maladaptive thoughts over the course of CT. Most often, it has been assessed by examining changes in specific depressogenic thinking patterns. Such changes have been measured using a variety of scales that often narrow in on one or more specific cognitive constructs. Examples of such measures include, but are not limited to, those that assess clients’ self-reported endorsement of pre-specified depressogenic thoughts and attitudes (Automatic Thoughts Questionnaire; Hollon & Kendall, 1980; Dysfunctional Attitudes Scale; Weissman & Beck, 1978) and client endorsement of specific depressogenic self-statements (Attributional Styles Questionnaire: Seligman, Abramson, Semmel, & von Baeyer, 1979). These self-report measures are typically given at a limited number of time points throughout treatment. One observer-rated measure of cognitive change that provides more flexibility in
assessing a wider variety of cognitive changes than those reviewed thus far is the Client Cognitive Change Scale (PCCS; Tang & DeRubeis, 1999). The PCCS has raters review session recordings or transcripts, noting instances of clients’ explicitly acknowledging any of seven categories of possible cognitive changes (e.g., bringing a schema into awareness, identifying a cognitive error). Thus, while many measures of specific self-reported depressogenic thinking exist, the PCCS is one of the only measures that allows for the more flexible assessment of broad cognitive change and skill acquisition. While the PCCS has several advantages, a key disadvantage is the time intensive coding required and reliance on naturalistic observation. In the present study, an original measure is used that aims to capture the flexibility the PCCS affords in a much briefer self-report format that can be quickly administered following each therapy session (see the In-session Cognitive Change Scale in the Methods section). As detailed below, this measure asks clients to reflect on the ways in which their thinking may have changed during the session that just occurred and the re-appraisal skills that may have fostered such change. Now that typical measures of assessing both the therapeutic alliance and cognitive change have been reviewed, in the below section I focus on the relations between these constructs and symptom change.

With regard to the therapeutic alliance, meta-analyses of the alliance-outcome literature show a small but reliable relationship between the alliance and outcome ($r$ ranging from .22-.28; Horvath, Del Re, Flückiger, & Symonds, 2011; Martin, Garske, & Katherine, 2000), with a stronger therapeutic alliance predicting above average symptom improvement. A recent meta-analysis also suggests a moderate association between the
therapeutic alliance and dropout (Sharf, Primavera, & Diener, 2010), such that clients with weaker alliances are more likely to drop out of treatment prematurely. In CT for depression, the smaller number of studies that have examined the alliance as a predictor of subsequent symptom change specifically have utilized mostly observer-ratings of the alliance and found less consistent evidence of an alliance-outcome relation (for discussion see Pfeifer & Strunk, 2015). Echoing this, a review by DeRubeis and colleagues (2014) highlights that when the results of the eight published alliance-outcome studies in which investigators accounted for temporal confounds are combined and weighted by sample size the resulting alliance-outcome estimate is considerably smaller ($r = .11$) than typical meta-analytic estimates (DeRubeis, Gelfand, German, Fournier, & Forand, 2014).

With regard to cognitive change predicting symptom reduction in CT, greater changes (i.e., reductions) in negative cognitions over the course of CT and within CT sessions have been found to predict significantly greater reductions in depressive symptoms post-treatment (Abel, Hayes, Henley, & Kuyken, 2016; McEvoy, Burgess, & Nathan, 2013), over between-session intervals (i.e., following sessions in which substantial cognitive change occurred, or “pre-gain” sessions; Tang & DeRubeis, 1999; Tang, DeRubeis, Beberman, & Pham, 2005), and over subsequent mid-treatment to post-treatment time periods (DeRubeis et al., 1990). It is important to note, however, that while most agree that cognitive change occurs over the course of successful treatment and relates to symptom change, skeptics of the causal role of cognitive change in CT argue that such change lacks specificity to CT and that evidence to support the temporal
precedence of cognitive change occurring prior to symptom reduction in CT is lacking (for reviews, see; Lorenzo-Luaces, German, & DeRubeis, 2015; Longmore & Worrell, 2007). Proponents of cognitive change as a potential causal mechanism of symptom reduction in CT acknowledge the lack of evidence supporting the specificity of cognitive change to CT, but suggest that this does not confirm or disconfirm whether cognitive change serves as a mechanism in CT or other treatments (Lorenzo-Luaces et al., 2015). These proponents further acknowledge that evidence of cognitive change predicting subsequent outcomes has been limited (particularly during acute treatment). They suggest that the lack of such evidence when combined with the abundance of evidence in support of the correlation between greater cognitive change and greater symptom reduction in treatment is consistent with the possibility that cognitive change could serve as a causal mechanism.

This study aims, in part, to explore the relation between baseline interpersonal vulnerability, the amount of cognitive change experienced early in treatment and the quality of the therapeutic alliance, and symptom change, in order to test the hypothesis that those with heightened vulnerabilities experience less cognitive change, form lower alliances, and thus, experience less symptom change. Overall, the above review suggests that CT processes of cognitive change, the therapeutic alliance, and interpersonal vulnerability have been found to predict treatment outcomes; however, given the mediational criteria outlined above, one must also consider the relationship between interpersonal vulnerability and these two CT processes.
The alliance and interpersonal vulnerability. A number of studies have highlighted the relationship of PD traits and related maladaptive attachment styles with the working alliance and therapeutic progress both in CT for depression as well as other treatments (for a discussion of the association between attachment styles and PD pathology see Meyer & Pilkonis, 2005). Clients with more adaptive, secure attachment styles and more positive attachment histories have been found to form stronger alliances with their therapists than those with maladaptive attachment styles (see meta-analytic review by Diener & Monroe, 2011; Mallinckrodt, 1991; Mallinckrodt, Coble, & Gantt, 1995). Additionally clients’ views of their interpersonal relations, which are often skewed when PD pathology is present (for a discussion see Luyten & Blatt, 2011), predict how clients rate the therapeutic relationship (Levin et al., 2012). Pathological personality characteristics, such as perfectionism, have also been found to predict a weaker formation of the therapeutic alliance (Taft, Murphy, Musser, & Remington, 2004; Zuroff et al., 2000). Often in conjunction with maladaptive personality traits, clients’ with a greater number of self-reported interpersonal problems pre-treatment have been found to form weaker working alliances across several cognitive treatments (Muran, Segal, Samstag, & Crawford, 1994; Saunders, 2001; Taft et al., 2004; Hardy et al., 2001; Renner et al., 2012). Replicating prior findings, a study by Holtforth and colleagues (2014) found that clients with overly hostile and hostile-submissive interpersonal styles (i.e., cold and highly introverted) reported significantly lower early therapeutic alliances than clients with overly friendly-submissive profiles (Holtforth et al., 2014). Taken together these
findings highlight the importance of considering the role of personality pathology and interpersonal distress when examining the alliance-outcome association.

**Cognitive change and interpersonal vulnerability.** While the above highlights the strength of the evidence supporting the relation between interpersonal vulnerability and the therapeutic alliance, less has been done to examine the relationship between interpersonal vulnerability and cognitive change in CT. In general, the relationship between interpersonal vulnerability and negative maladaptive beliefs, thoughts, and attitudes has been well documented (Fournier, DeRubeis, & Beck, 2012; for a review see Hames et al., 2013; Hopwood, Schade, Krueger, Wright, & Markon, 2012). In the context of CT for depression, clients’ interpersonal vulnerability has been related to maladaptive cognitive beliefs measured at the start treatment (McEvoy, Burgess, & Nathan, 2013), with clients higher in interpersonal vulnerability pre-treatment reporting more negative cognitions. Likewise, in a study of non-treatment seeking adults, Roberts and colleagues found that insecure adult attachment styles, often predictive of interpersonal problems, were associated with dysfunctional attitudes which predicted lower levels of self-esteem and increases in depressive symptoms over time (Roberts, Gotlib, & Kassel, 1996). To my knowledge, however, no studies have assessed the predictive relation of interpersonal vulnerability at baseline on subsequent cognitive change over the course of CT for depression. This study aims, in part, to explore this association.

Taken together the above sections provide evidence to: (1) support the relation between negative maladaptive thoughts, interpersonal vulnerability, the working alliance and depressive symptoms in CT; (2) support the relation between interpersonal
vulnerability and the alliance in CT, and; (3) raise the possibility that interpersonal vulnerability may be related to cognitive change occurring over the course of CT for depression while also highlighting the need to better understand the impact of interpersonal vulnerability on cognitive change. Below is a review of the few studies that have more specifically examined either the alliance or cognitive change as mediators of the relation between interpersonal vulnerability and outcome in CT.

**Cognitive change and the alliance as mediators: Preliminary evidence.** The above sections support a relation between interpersonal vulnerability and the therapeutic alliance, cognitive change, and ultimately treatment outcome, raising the possibility that interpersonal vulnerability may affect CT outcomes indirectly by a relationship mediated by CT processes. Specifically, the relationship between interpersonal vulnerability and CT outcomes may be mediated by the alliance for example, such that, increased interpersonal vulnerability predicts decreased alliance scores, which in turn predict less marked symptom improvement (McEvoy, Burgess, & Nathan, 2014). Put another way, the alliance may account for the relationship between interpersonal vulnerability and outcome in CT. McEvoy and colleagues (2014) examined this possibility in a community clinical sample receiving either individual (N = 84) or group (N = 115) CBT for anxiety or depression. They found that heightened interpersonal problems at baseline were associated with increased risk of dropout and less symptom improvement in group, but not individual, CT. No evidence of a mediational relationship between interpersonal problems, the alliance, and outcomes were found in either treatment type; however, a separate set of analyses on the same sample revealed that the correlation between IIP and
BDI-II (Beck Depression Inventory - 2nd Edition; Beck, Steer, & Brown, 1996) pre-to-post treatment change scores in CBGT was accounted for, in part, by changes in clients’ negative cognitions throughout treatment, such that the association between IIP and BDI-II change was no longer significant when cognitive change scores were accounted for (McEvoy, Burgess, & Nathan, 2013).

In contrast, Hardy and colleagues did find support for a mediational relation between interpersonal vulnerability, the alliance, and outcome in a smaller (N = 24) sample of depressed clients undergoing individual CT. Specifically, they found that the relationship between clients’ pre-treatment interpersonal vulnerabilities and post-treatment depressive symptoms was mediated by the therapeutic alliance, such that clients’ who were self-reportedly avoidant and socially under-involved had greater difficulty forming strong alliances with their therapists which in turn predicted suboptimal post-treatment depression outcomes (Hardy et al., 2001). Howard and colleagues (2006) also found evidence of a mediational relationship in a small sample of multiple sclerosis clients undergoing 16 weeks of CT for depression. Specifically, clients’ IIP scores at baseline significantly predicted post-treatment BDI-II scores and week 4 client-reported alliance; however, when controlling for week 4 alliance scores the relationship between the IIP and post-treatment BDI-II scores was no longer significant, supporting the alliance as a mediator of the relation between interpersonal vulnerability and post-treatment outcomes (Howard, Turner, Olkin, & Mohr, 2006).

Taken together, the above sections provide support for the following: (1) clients’ self-reported interpersonal vulnerabilities (i.e., interpersonal distress, skill deficits, and
maladaptive personality traits) predict symptom reduction and dropout in treatment for depression; (2) clients’ self-reported interpersonal vulnerabilities are related to the formation of the therapeutic alliance in CT; (3) clients with self-reported interpersonal vulnerability tend to also have more negative maladaptive beliefs, thoughts and attitudes; (4) clients’ self-reported interpersonal vulnerability may be related to the amount of cognitive change experienced in CT for depression, and; (5) both cognitive change and the therapeutic alliance have been related to outcome in CT and may serve as potential mediators of the association between interpersonal vulnerability and outcome.

It is important to note that, with the exception of one, all of the above studies rely exclusively on clients’ self-reported interpersonal vulnerability. As we will review below, most studies assess interpersonal distress and maladaptive interpersonal styles using some variant of the IIP (a self-report measure). Similarly, the vast majority of the above research examining clients’ pathological personality characteristics utilizes either self-reported client traits or diagnostic information regarding PDs. While frequently used self-report measures offer reasonable reliability and may aide replication efforts\(^1\), the exclusive reliance on self-report information precludes an understanding of how others’ (i.e., acquaintances, strangers, close loved ones, and therapists) perceptions of the clients’ interpersonal vulnerabilities might affect their depressive symptoms and treatment progress. The below and final section of this introduction reviews some of the most

\(^1\) Replication efforts are only aided if the same version of a given self-report measure is utilized across studies and other cross-study differences are held constant. It is important to note that while there is often some consistency in the diagnostic approach to assessing PDs, interpersonal distress and styles are often assessed with one of the many different version of the IIP. As reviewed below, although these versions have some common items, differences in scoring approaches make comparisons between studies using different IIP derivatives difficult (McEvoy, Burgess, Page, Nathan, & Fursland, 2013)
commonly used self-report measures of interpersonal vulnerability and related constructs, expands upon the implications of relying exclusively on self-report measures of interpersonal vulnerability, and poses an alternative approach to assessing such constructs that may offer increased ecological and predictive validity.

Ways of Measuring Interpersonal Vulnerability and Associated Traits

Given the well-established association between interpersonal vulnerability and psychological difficulties, several empirical efforts have focused on designing self-report and diagnostic measures that best capture interpersonal vulnerabilities and associated maladaptive personality traits.

Common self-report and assessor-administered measures. With regard to assessing PDs, PD traits, and pathological variants of personality a combination of self-report and assessor-administered interviews are most typically used. PD diagnoses are most commonly assessed using the symptom criteria laid out in the Diagnostic and Statistical Manual of Mental Disorders (DSM; for most recent version, DSM-5, see American Psychiatric Association, 2013). These symptoms may be assessed by an interviewer either informally or more formally using, for example, the Structured Clinical Interview for DSM-IV Axis II Personality Disorders (SCID-II; First, Gibbon, Spitzer, Williams, & Benjamin, 1997). In the more formal SCID-II approach, clinician or evaluator has the client complete the self-report SCID-II pre-screener questionnaire, in which the client reports on various problematic personality features. The evaluator then uses the structured interview questions to follow-up with the client on screener items he/she endorsed. The structured interview questions probe formal DSM criterion for each
PD by asking the client a series of questions. Based on the client’s answers, the evaluator then assigns scores between 1 (not present) and 3 (definitely present) for the given criteria. The evaluator then adds up the number of “definitely present” scores for each PD and compares this quantity to the respective clinical threshold to determine if a formal diagnosis is warranted. While the number and nature of symptoms required for the various PDs has shifted slightly over the years, the same basic format for assessing PDs remains, that is, diagnoses are determined based on clients’ endorsements of questions reflecting common symptoms, which are then determined, by the interviewer, to fit with the pre-specified criteria. It is important to note that the above means of assessing PDs require the client to accurately identify and report on their PD symptoms when prompted and only taps into the value of the assessor’s perspective to the extent that the client offers such information.

The sufficiency and validity of using a diagnostic approach akin to that described in the DSM, which calls for clinicians to make dichotomous decisions regarding personality pathology, has been widely debated and called into question in recent years (for a discussion see, Hopwood et al., 2012). One of the common complaints levied against prior versions of the DSM is the diagnostic comorbidity and heterogeneity observed among PDs. While Section II of the DSM-5 details diagnostic criteria required to make a diagnosis in an approach very similar to that described above in prior DSM editions, the supplemental Section III of the DSM-5 aims to characterize personality disorders as more dimensional impairments in personality functioning and traits. Specifically, the Personality Inventory for DSM-5 (PID-5; Krueger, Derringer, Markon,
Watson, & Skodol, 2012) is a 220-item self-report inventory which indexes the DSM-5 personality domains and their respective facets. Domains are designed to capture pathological variants of the five-factor model of normative personality (FFM; McCrae & Costa, 1997). Specifically the PID-5 assesses traits of Negative affectivity (like FFM neuroticism), Detachment (like low extraversion), Antagonism (like low agreeableness), Disinhibition (like low conscientiousness), and Psychoticism (like Openness to Experience). Initial findings suggest that this trait-based DSM-5 model captures reliable variance in PDs and interpersonal dysfunction common in personality pathology, and has been posed as a potential solution to some of the empirical problems with dichotomous PD diagnoses of prior DSM editions (Hopwood, Thomas, Markon, C, & Krueger, 2012). While more dimensional approaches to defining personality pathology appear promising, all of the above methods rely, to an extent, on clients’ reporting of such traits.

Lastly, non-pathological personality traits, or the traits laid out by the FFM, may also be important to consider when assessing interpersonal dysfunction. A series of structural analyses of trait descriptors in the natural language have consistently revealed the FFM’s five broad factors: extraversion, agreeableness, conscientiousness, neuroticism and openness to experience (John & Srivastava, 1999). Variants of these FFM traits have been correlated with several maladaptive behaviors that may be precursors or correlates to interpersonal dysfunction including but not limited to: substance use, unhealthy eating, risky driving and sex, suicide, conflict, criminality, and clinical disorders; as reviewed by Lengel, Helle, DeShong, Meyer, & Mullins-Sweatt, 2016). These FFM traits are almost exclusively assessed using self-report measures, with a number of measures available but
perhaps the most popular being the NEO Personality Inventory (NEO PI-R; Costa & McCrae, 1992). Thus, no shortage of self-report based measures exists for assessing personality traits associated with interpersonal dysfunction. As reviewed below, interpersonal distress and interpersonal styles more specifically are also typically assessed using self-report formats.

**Self-report interpersonal vulnerability: IIP-32 and its predecessors.** Given the importance of understanding the role of interpersonal distress in relation to psychological wellbeing, Horowitz and colleagues developed the Inventory of Interpersonal Problems in the late 80s (IIP-127; Horowitz et al., 1988). Since this time, several different versions of the IIP have been generated that vary the number of items and scoring of the original IIP-127. A brief review of the various iterations of the IIP scale is provided below along with details and support for the version of the IIP (IIP-32; Barkham, Hardy, & Startup, 1996) used in the current study.

The original IIP-127 consists of 127 items divided into two major types: items beginning with ‘I find it hard to…’ and items beginning with ‘I do too much…’ The original factor-analysis of the IIP-127 identified six scales but suffered from serious limitations including an insufficient sample size and an unrepresentative sample comprised mostly of females (McEvoy, Burgess, Page, Nathan, & Fursland, 2013). A subsequent factor analysis addressing these limitations identified the eight factor structure more commonly used today: Too aggressive, Too open, Too caring, Too dependent, Hard to be assertive, Hard to be sociable, Hard to be supportive, and Hard to be involved (Barkham, Hardy, & Startup, 1994). Today, ten derivatives of the original IIP-127,
including the version utilized in this study, have been published and are more commonly used across a range of empirical studies than the IIP-127 (for a review see; Hughes & Barkham, 2005).

These ten IIP derivatives can be grouped into one of two different versions based on the analytic approach they employ to score and interpret responses: versions that utilize the circumplex model of personality (Leary, 1957; for an example see the IIP-C; Alden, Wiggins, & Pincus, 1990; Horowitz, Alden, Wiggins, & Pincus, 2000) and versions that utilize a factor-analytic based approach (i.e., IIP-127 and the more recent IIP-32; Barkham et al., 1996; Horowitz et al., 1988). In contrast to factor-analytically derived subscales, the Circumplex derived scales are based on a taxonomical model of interpersonal behaviors formed by placing different behaviors in a circle with respect to two orthogonal dimensions: dominant vs. submissive and hostile vs. friendly (as reviewed by Hughes & Barkham, 2005). One difficulty with using the circumplex approach is that in clinical settings, it is important that client’s self-report questionnaires be brief and readily interpretable to clinicians. Creating and interpreting an IIP profile using circumplex analysis is much more challenging and time-consuming than calculating scale scores as is done with factor-analytically derived versions. Thus, it has been suggested that the shorter factor-analytic based versions of the original IIP-127 may be preferred over their circumplex counterparts in clinical settings (Hughes & Barkham, 2005; McEvoy, Burgess, Page, Nathan, & Fursland, 2013). The IIP-32 (Barkham et al., 1996) was developed to provide a readily interpretable and briefer version of the IIP-127 which could be of greater utility in clinical settings. Given the brevity and established
psychometric properties detailed below, the IIP-32 was chosen as the self-report assessment of interpersonal problems in the current study.

The IIP-32 contains 32 self-report items selected from the original IIP-127 scale in order to represent the eight factor-analytically (see Barkham et al., 1994) derived subscales identified in the original version. Barkham and colleagues (1996) reviewed the empirical work conducted to evaluate the creation of the IIP-32 and suggested that the IIP-32 sacrifices little compared to the original IIP-127 in terms of psychometric properties. In an attempt to replicate the factor structure identified by Barkham et al. (1996) and validate the IIP-32 against other measures, McEvoy and colleagues (2013) examined the psychometric properties, factor variance, and external validity of the IIP-32 (Barkham et al., 1996) in two treatment-seeking samples with primary anxiety and depressive disorders (n = 504; McEvoy, Burgess, Page, Nathan, & Fursland, 2013). Confirmatory factor analyses showed that the previously established eight-factor structure of the IIP-127 provided the best fit. The IIP-32 also demonstrated evidence of external validity against other well-validated measures of anxiety, depression, and quality of life. Specifically, a tendency to subjugate one’s needs by prioritizing others’ needs and difficulties being sociable were consistently associated with a poorer quality of life, and more severe anxiety, depression, and eating disorder symptoms. Additionally, all subscales showed a significant positive correlation with the BDI-II, such that heightened interpersonal vulnerability was related to more severe depressive symptoms.
Over-Reliance on Self-Reported Interpersonal Vulnerability

While self-report and diagnostic measures of interpersonal distress and related PD traits are desirable in that they often exhibit satisfactory reliability and are commonly used, a multi-modal approach to assessing such constructs may foster a more ecologically valid and complete understanding. More specifically, an over-reliance on self-report measures of interpersonal vulnerability may be particularly problematic in that such measures are based on the questionable assumption that people are able to accurately and comprehensively report on deficits in their interpersonal skills and a wide variety of pathological personality traits (for a discussion see Leising, Krause, Köhler, Hinsen, & Clifton, 2011; Hogan, 1996; Mulder, 2002). However, this may not be the case as self-report responses are susceptible to biases including self-deception, limited insight, and socially desirable responding (Sasso & Strunk, 2013). While scales have been created in order to assess such biases, they have yielded disappointing results in detecting and correcting for such distortions (Connelly & Ones, 2010; Ellingson, Sackett, & Hough, 1999). Indeed, self-other agreement in judging personality pathology and interpersonal functioning is generally weak (e.g., Clifton, Turkheimer, & Oltmanns, 2005; Klonsky & Oltmanns, 2002; Leising et al., 2011; Oltmanns & Turkheimer, 2006). Additionally, self-reports of interpersonal functioning may be further limited to the extent that one views his/her interpersonal problems as adaptive (e.g., as can be the case with avoidance).

With regard to the predictive validity of interpersonal vulnerability and related constructs, how one is perceived by others may be more consequential than one’s self-views. For example, in a recent meta-analysis others’ ratings (e.g., family, friends, or
even complete strangers) of an individual’s personality had substantially greater validity in predicting academic achievement and job performance than self-report ratings (Connelly & Ones, 2010). Another study found that while both clients’ and informants’ reports of PD diagnoses and dimensional PD scores independently predicted subsequent depressive symptoms, only informants’ reports predicted subsequent social adjustment (Klein, 2003). Finally, relative to self-reports, peer reports of PD traits were better predictors of early discharge from the military (Oltmanns & Turkheimer, 2009). These discrepancies in the predictive validity of self- vs. other-informants of interpersonal vulnerability raise the possibility that relying exclusively on self-reports of such constructs when examining their effects in a therapeutic setting may impede the identification of true relations between clients’ traits and treatment outcome.

**Means of Assessing Others’ Perceptions of Interpersonal Vulnerability**

Capturing quick first impressions of clients’ traits may serve as an effective means of assessing others’ perceptions, as research suggests that interpersonal difficulties may be detected in rapid judgments made with little to no conscious effort (Hammen, 2005; Oltmanns, Friedman, Fiedler, & Turkheimer, 2004). In fact, findings show that one’s pathological personality traits may be recognized by strangers who watch very brief (≤ 5 min) video-clips of that individual, or “thin slice” clips. For example, Fowler, Lilienfeld, and Patrick (2009) found that thin slice ratings of psychopathy features in maximum-security inmates were related to well-validated assessments of psychopathy, with ratings of violence proneness and antisocial PD correlating significantly with the number of evaluator determined antisocial PD symptoms. Oltmanns and colleagues
(2004) found that untrained undergraduate students were able to make reliable judgments about target individuals’ PD related personality traits and non-pathological five factor model traits (Costa & McCrae, 1992) on the basis of 30 second thin slice clips taken from diagnostic interviews. Expanding on this work, Friedman and colleagues (2007) found that untrained raters were able to reliably predict scores on personality pathology measures obtained from target individuals’ close peers, on the basis of thin slice clips of target individuals describing themselves (Friedman, Oltmanns, & Turkheimer, 2007). The above findings suggest that quick judgments regarding one’s personality traits mirror meaningful information about how such traits are perceived by others, including traits in the clinical range that are often associated with interpersonal difficulties.

Based on this notion, Sasso and Strunk (2013) assessed the utility of thin slice ratings of personality traits in an adult clinical population. In a sample of 66 clients participating in a naturalistic study of CT for depression, they had trained raters assess clients’ personality characteristics and PD traits on the basis of thin-slice video clips taken from clients’ pre-treatment intake assessments. The five higher-order FFM traits, PD characteristics shown to be discernable to thin slice raters, and likeability were assessed using the Thin Slice (TS) observer ratings scale, which Sasso and Strunk (2013) created by adopting items used to assess these constructs in prior thin slice studies (see Friedman et al., 2007; Oltmanns et al., 2004). The TS scale also contained three items to assess clients’ interest in therapy. They found that TS ratings of clients’ perceived therapy interest and neuroticism predicted lower odds of treatment dropout. Ratings indicative of heightened client extraversion predicted greater symptom change across
treatment; whereas, ratings indicative of heightened Avoidant- and Schizoid-like PD traits predicted less marked symptom improvements. Raters’ perceptions of heightened client agreeableness and likeability also showed relations to greater symptom reductions. Interestingly, clients’ self-reported PD traits and diagnostically determined PD status failed to account for the above identified thin slice-outcome relations. These findings support the clinical utility of quick, thin slice impressions of client traits; as such ratings may offer the opportunity to pre-emptively identify clients at heightened risk of dropout and non-response.

The above findings establish the potential in assessing personality traits, frequently linked to interpersonal vulnerability, using a thin slice methodology; however, observer ratings of clients’ behavior in interpersonal contexts may serve as an important compliment to these thin slice ratings, particularly when clinical outcomes are of primary interest. Specifically, all of the above studies obtain observer ratings on the basis of video-recordings of the target individual describing themselves in some capacity (i.e., via answering diagnostic questions, describing their personality, or discussing the important people in their life). Thus, ratings may reflect the target individual’s self-image more than their actual interpersonal skills or behaviors. While one’s self-image has clear implications on interpersonal vulnerability, ratings of clients’ interpersonal functioning based on interactions in which the client is not explicitly describing themselves but rather asked to perform an interpersonal behavior, may also be important to understanding how clients’ interpersonal vulnerability is perceived by others.
**Performance on interpersonal role plays.** Prior research on the behavioral tendencies of individuals with depression in interpersonal interactions provides support for the notion that additional information might be garnered by having observers provide ratings in such a context. Individuals with depression have been found to initiate as few as half the number of interpersonal actions, including starting and carrying on conversations (Hames et al., 2013), and when they do engage in such interactions, they tend to introduce more negatively themed content including self-disclosure of negative events, feelings, and self-evaluations at potentially inappropriate times (Segrin, 2000). These tendencies, which leave those with depression vulnerable to social rejection and isolation, may be readily noticed by objective observers but potentially uncaptured via traditional self-report methods. Additionally, while one might expect thin slice ratings of personality to be strongly related to performance ratings in interpersonal contexts, it is possible that the latter may provide a more ecologically valid proxy of interpersonal vulnerability. Indeed, much of the research on “social skills” and “social competence” in non-clinical settings tends to utilize skill ratings from people other than the target person (i.e., teachers, parents) as well as observer-ratings of one’s behavior in interpersonal role-plays.

With this in mind, Leising and colleagues (2011) set out to create an interpersonal role-play with a series of tasks that would expand upon those used in the social skills training literature by containing interactions which may be particularly influenced by pathological personality traits, and thus better suited to capture interpersonal dysfunction.
(i.e., “Assert a claim”; see Leising et al., 2011). Their “Standardized Interaction Task” or SIT contains 17 brief (≤ 90s) interpersonal roleplays, each with simple instructions regarding an interpersonal situation the target and researcher are to engage in and a behavior the target person is expected to show. Leising and colleagues (2011) assessed interpersonal vulnerability and pathological personality using three different modalities (self-report questionnaires, diagnostic interviews, and the SIT) within the same sample of non-treatment seeking adults (n = 100). They found that clients’ self-reported interpersonal difficulties (IIP-64 scores) largely mirrored their own judgments of their performance on the SIT (i.e., self-report SIT ratings), with those reporting higher levels of self-attributed interpersonal distress also being more skeptical of their role-play effectiveness. In contrast, they found an overwhelming lack of overlap across the different informants (i.e., agreement between self- and observer-ratings on mean SIT performance across all tasks $r = .02$). Similarly, the correlation between individuals’ IIP-64 total scores (i.e., general self-reported interpersonal problems) and SIT-observer ratings was non-significant ($r = .04$). Taken together, these findings raise the possibility that self-reported interpersonal functioning may fail to capture behavioral variance relevant to others’ perceptions of such functioning and may not reflect the reality of ones’ interpersonal problems as perceived by outsiders.

---

2 A slightly modified version of this standardized interaction task (SIT) is used in the current study.
3 IIP-64 (Horowitz et al., 2000), SCID-II (First, Gibbon, & Spitzer, 1997) and the SIT.
4 Leising et al. (2011) excluded participants who were currently in psychiatric treatment or using psychotropic medications. Of the 100 included participants, 7 individuals received categorical PD diagnoses following SCID-II assessment: 4 obsessive-compulsive, 2 schizoid, and 1 borderline diagnosis.
Purpose of This Study

This study proposes to address, in part, the question of what predicts symptom change in CT for depression by examining the role of interpersonal vulnerability, particularly as perceived by others. As reviewed above, a scarcity of research has examined the clinical utility of assessing clients’ interpersonal vulnerability from multiple perspectives. Interpersonal vulnerabilities are heightened in those with depression and readily noticeable to others in a manner that may be discordant with how the individual themselves views those vulnerabilities. Additionally, self-reported interpersonal problems and related maladaptive personality traits have been found to predict significant variability in symptom change, dropout, and mechanisms of change in treatments for depression and more specifically CT (see above review including findings from, but not limited to; Diener & Monroe, 2011; Hardy et al., 2001; Newton-Howes et al., 2006; McEvoy, Burgess, & Nathan, 2013; Renner et al., 2012; Swift & Greenberg, 2012). While little work has been done to examine the predictive validity of observer rated interpersonal vulnerability in a treatment context, how clients’ are perceived by others may explain meaningful variation in their therapeutic experience. The TS and SIT observer measures may provide ecologically valid measures of clients’ interpersonal vulnerabilities as perceived by others, and the TS measure has shown predictive validity above and beyond commonly-used self-report measures in predicting symptom change and dropout in CT for depression specifically (Sasso & Strunk, 2013).

This study aims to build upon the above literature and provide a more comprehensive view of how clients’ interpersonal problems and traits may be impacting
CT outcomes by utilizing a multi-method approach to assess interpersonal vulnerability in depressed clients. There are three main objectives in this study: (1) examine how clients’ interpersonal vulnerabilities predict treatment outcomes (i.e., symptom change and dropout) in CT for depression; (2) evaluate how clients’ interpersonal vulnerabilities, as perceived by others, are related to two key CT process measures (cognitive change and the alliance), and; (3) examine whether these process measures account for the relation between interpersonal vulnerabilities and outcome.

Hypotheses

In evaluating these three main objectives, I propose the following primary hypotheses:

Objective 1: Hypotheses 1A-1B and 2A-2B. A primary aim of this study is to determine the relationship between perceptions of clients’ interpersonal vulnerabilities and subsequent treatment outcomes including changes in depressive symptoms and attrition (i.e., dropout). I predict that clients’ SIT- and TS-observer ratings, made based on video clips taken from clients’ intake evaluations, will be significant predictors of subsequent slope of depressive symptoms over the course of acute treatment and show numerically stronger effects on symptom change than self-reported measures of interpersonal vulnerability (Hypotheses 1A). I also predict that the same pattern of results will hold when examining both observer rated and self-report interpersonal vulnerability as predictors of risk of early attrition (i.e., dropout; Hypotheses 1B). I will also examine the extent to which a therapist’s perspective of clients’ interpersonal vulnerabilities
predicts outcome. I will explore the effects of TS-therapist ratings\(^5\) on subsequent symptom change and dropout. My expectation is that the TS subscales\(^6\) identified as predictors when observer ratings were examined will also be significant predictors of symptom change (Hypothesis 2A) and dropout (Hypothesis 2B) when assessed by therapists. I will also test whether relations between TS subscales and outcome vary as a function of whether the rater was a therapist or an observer, by examining interactions between TS subscales of interest and “rater type”. Insofar as therapists’ unique perspectives have important consequences for the process of change in CT, I would expect therapist ratings to predict more strongly than observer ratings.

Objective 2: Hypotheses 3-4. A second aim of this study is to examine how others’ perceptions of clients’ interpersonal vulnerabilities relate to two key process measures; in-session cognitive changes and the therapeutic alliance. I predict that heightened interpersonal vulnerability, as assessed by TS- and SIT-observer ratings at clients’ intake evaluations, will significantly predict lower average therapeutic alliances (Hypotheses 3) and lower average in-session cognitive change scores, early in treatment (Hypotheses 4). To explore the extent to which a therapist’s perspective of clients’ interpersonal vulnerability predicts alliance and cognitive change scores, I will examine the effects of TS-therapist ratings on average alliance and cognitive change scores early

---

\(^5\) As described below, TS-therapist ratings are made by each client’s therapist prior to the start of treatment with that client and solely on the basis of the details provided by the client in their ≤ 5-minute intake thin-slice video clips.

\(^6\) As reviewed below, exploratory factor analyses revealed a slight difference in the factor structure for therapist vs. observer TS ratings, with therapists’ being better able to distinguish histrionic traits than observers (i.e., a 4-factor structure for TS-therapist vs a 3-factor structure for TS-observer). Therefore, when I specify the TS subscales that will predict for both therapist and observer informants, I am referring specifically to the identical subscales of Pro-treatment behaviors and Neuroticism, and also the near identical Sociability subscale, as described below.
in treatment. Additional exploratory analyses will examine how the above relations vary across observer and therapist TS informants and explore relations of the above constructs to slope of alliance and in-session cognitive change across early sessions.

**Objective 3.** The final aim of this study will be to examine the possibility that clients’ average therapeutic alliance and self-reported in-session cognitive change scores observed early in treatment account for any significant relations identified between interpersonal vulnerability at intake and subsequent slope of symptom change. For the purpose of these analyses the variables reflecting interpersonal vulnerability will be examined in separate models respectively and analyses will remain exploratory.
Chapter 2: Methods

Data Collection and Cleaning

All research for this study was conducted in the Depression Treatment and Research Clinic located at the Ohio State University (OSU). The OSU institutional review board reviewed and approved all research activities associated with this study prior to the beginning of data collection.

Participants

The sample was comprised of 126 adults with a primary diagnosis of MDD who participated in an out-patient treatment study examining determinants of outcome in CT for depression. Treatment was provided free of charge to clients participating in the study. Recruitment was conducted using advertisements and word-of-mouth in the central Ohio region. Interested individuals were directed to contact researchers for a pre-screening phone interview. Upon first contacting study staff, participants completed a brief phone screen. Based on the phone screen, those deemed appropriate for the study were scheduled for an initial intake assessment. Criteria assessed via phone screen were: likely to meet criteria for current major depressive disorder, no self-reported history of manic episodes or current substance dependence, stable on any anti-depressant medication (ADM) or not on any ADM (with no plans to initiate ADM treatment during the course of the study), and willing to commit to 16 weeks of treatment. At the
subsequent initial intake assessment, full inclusion / exclusion criteria (as described below) were assessed. Those who met these criteria were invited to participate in the study and receive 16 weeks of CT.

Inclusion criteria for participation in the treatment study were: (a) diagnosis of MDD, according to DSM-IV criteria (American Psychiatric Association, 1994); (b) 18 years or older, and; (c) willing and able to give informed consent. Exclusion criteria were: (a) current Axis I disorder other than MDD if it constituted the predominant aspects of the clinical presentation and if it required treatment other than that being offered; (b) Bipolar I disorder or psychosis; (c) subnormal intellectual potential (IQ < 80); (d) evidence of any medical disorder or condition (including pregnancy) that could cause depression; (e) clear indication of secondary gain (e.g. court ordered treatment), and; (f) current suicide risk sufficient to prevent treatment on an out-patient basis. In addition to meeting these criteria, clients on anti-depressant medication were asked to maintain a stable dosage over the course of the study. All participants provided written informed consent prior to all research activity.

Out of 193 clients invited to partake in the intake assessment, 43 did not present for the scheduled appointment and 1 completed a portion of the assessment but opted out prior to assessment of eligibility. Of the 149 clients who participated in the intake assessment, 23 did not qualify for the study. Criteria for excluding these clients included symptoms deemed to be insufficient for a major depressive episode diagnosis (n = 8), history of Bipolar or Manic Episodes (n = 6), depression did not constitute the primary
diagnosis ($n = 4$), substance dependence within the last 6 months ($n = 3$), presence of a psychotic disorder ($n = 1$), and unwilling to end current therapy ($n = 1$).

**Demographics and descriptive statistics.** The sample of 126 clients who met inclusion criteria and enrolled in the study was comprised of 52 (41%) males and 74 (59%) females. The mean age of clients was 32 years ($SD = 13.06$, range 18-70). The ethnic composition of the sample was 82% Caucasian ($N = 103$), 7% African-American ($N = 9$), 8% Asian ($N = 10$), 2% Hispanic ($N = 3$), and 1% American Indian ($N = 1$). The majority of the sample (59%) reported having never been married, whereas 23% reported being currently married or cohabitating. The majority (72%) of the sample had completed either a 4-year college degree ($N = 30$), part of college ($N = 53$), or a 2-year college degree ($N = 7$).

**Diagnoses and symptoms at intake.** All of the 126 clients met criteria for current MDD as assessed by the SCID-I at intake. At this assessment point, clients were found to have a mean BDI-II score of 32.80 ($SD = 8.81$, range 11-56), indicating that on average clients were experiencing severe depressive symptoms. Table 1 summarizes diagnostic specifiers and common comorbid Axis I diagnoses across clients. Thirty-seven clients had no current comorbid disorders, but the majority of clients had at least one comorbid Axis I diagnosis. The number of concurrent diagnoses ranged from one to five. Anxiety disorders were the most common comorbidity, and among these, Social Phobia was reported most frequently ($n = 54$). Seventy-one clients reported recurrent depressive episodes (ranging from two to five previous episodes).
Assessors and Therapists

Five advanced graduate students served as both study assessors and therapists. The therapist to whom a client was assigned never conducted that client’s assessments. Clients were quasi-randomly assigned to therapists and assessors (given the constraint that the client’s assessor could not be the client’s therapist and vice versa). Therapists received approximately 60 hours of clinical training in CT, with a focus on experiential learning through role-play. Therapists received weekly individual and group supervision over the course of treatment from principal investigator, Daniel R. Strunk, Ph.D. Therapists followed procedures described in the treatment manual (Beck et al., 1979).

Session Frequency

Individual sessions were scheduled to take place twice per week for the first four weeks for all clients. After the fourth week of treatment, the majority of clients decreased session frequency to one session per week, but clients were offered the option of continuing to schedule twice weekly sessions up to week 12 of treatment. Following week 12 of treatment, all clients had sessions only once per week. The mean number of sessions that clients received in the first four weeks of treatment was 6.12 (SD = 1.51, range 1-8) and the mean number of sessions clients received in the first 12 weeks of treatment was 13.13 (SD = 3.88, range 1-24). Over the full course of treatment the mean number of sessions received per client was 15.93 (SD = 5.14, range 1-28). These
represent conservative estimates of clients’ session frequency as such estimates include those clients who were deemed treatment dropouts, as defined below.\(^7\)

**Measures**

**Diagnoses.** At the intake evaluation, Axis I diagnoses were assessed with the Structured Clinical Interview for the DSM-IV (SCID-I; First, Spitzer, Miriam, & Williams, 2002).

**Outcome measures.**

**Dropout.** Clients were deemed dropouts if they prematurely terminated treatment prior to receiving an adequate dose (i.e., terminating four weeks or more prior to their post-treatment assessment date which equates to a failure to complete 12 full weeks of treatment). Clients who missed sessions without declaring an intention to discontinue treatment were encouraged to return to therapy. If after four weeks of repeated attempts a given client’s therapist had not been able to contact his/her client, the client was deemed a dropout. If the client returned to treatment at any time within this four week period, they were not considered a dropout. Of the 126 clients, 17% (\(N = 21\)) were determined to be dropouts. Of these 21 dropouts, 1 client did not attend his/her first scheduled session and 2 clients only attended session 1 of treatment. 43% (\(N = 6\)) dropped out early in treatment (i.e., prior to week 4, but after completing at least 2 sessions) and 57% (\(N = 12\)) dropped out in that latter half of treatment (at or following the week 4 mark).

**Depressive symptoms.** The main outcome variable in this study, depressive symptoms, was assessed using the BDI-II. The BDI-II is a 21-item self-report instrument

---

\(^7\) The mean number of sessions per client, when those deemed dropouts were excluded, was: 6.44 over the first four weeks of treatment, 14.30 across the first 12 weeks of treatment, and 17.77 over the full course of treatment.
used to assess the severity of symptoms of depression. Respondents are asked to describe the way they have been feeling during the past week\(^8\) by rating each item (e.g., sadness, pessimism, loss of pleasure) on a scale from 0 to 3. Thus, possible scores are 0 (minimal depression) to 63 (high depression). This scale was administered at each therapy session and at clients’ intake evaluation. This measure has repeatedly demonstrated excellent internal consistency, with coefficient alphas averaging .91 or higher (Dozois, Dobson, & Ahnberg, 1998; Steer, Kumar, Ranieri, & Beck, 1998). Estimates of the convergent validity of the BDI-II scores have produced moderate to large correlations with alternative measures of depression severity (Beck et al., 1996).

**Therapeutic process measures.**

**Therapeutic alliance.** The therapeutic alliance was assessed using the short form of the client-rated Working Alliance Inventory (WAI-C-S; Horvath & Greenberg, 1989; Tracey & Kokotovic, 1989), which consists of 12 items evaluated on a 7-point Likert scale (0 = never, 6 = always). The WAI is a commonly used measure of the quality of the therapeutic alliance and has shown satisfactory reliability. In this study, clients responded to the WAI immediately following the first five therapy sessions. Clients completed the form with the assurance that their responses would not be seen by the therapist and were asked to seal their form in an envelope for study personnel other than the therapist to open and process. Across sessions 1 through 5, Cronbach’s alphas for client ratings

---

\(^8\) Modified instructions for the BDI-II were used to inquire about the past week, instead of the past 2 weeks, in order to assess the degree of symptom change on a session-to-session basis (for a similar modification, see Dimidjian et al., 2006)
ranged from .93 to .98 for the WAI-C total sum score, which will be used in all analyses below.

**In-session cognitive change (CC).** In-session cognitive change was assessed using an original 5-item measure. The measure asks clients to focus on the therapy session that just ended and indicate the extent to which they agree with the five items on a 7-point Likert scale (0 = Not at all, 6 = completely). Included items are designed to capture the extent to which clients re-evaluated negative automatic thoughts or core beliefs within the session that just occurred (i.e., “I took note of what I was thinking and worked to develop a more balanced view”). As reviewed above, this original measure was used as no existing measures flexibly captures broad in-session cognitive changes in a brief and easy to administer format. The CC was administered immediately following the first five therapy sessions. As this measure was newly developed for use in this study, I evaluate its reliability and validity below and report on these findings in the results section.

**Predictors of interest.**

**Inventory of interpersonal problems (IIP-32).** As reviewed in some detail above, the IIP-32 (Barkham et al., 1996) is a relatively brief 32-item self-report measure which yields a total score that reflects overall interpersonal difficulty and different types of interpersonal problems respectively. Respondents are asked to review a list of problems that people report in relating to others, including things individuals find hard to do and things individuals may do too much of. For each item, respondents then indicate how distressing the particular problem has been for them on a 5-point Likert-type scale.
ranging from 0 = not at all to 4 = extremely. The IIP-32 has demonstrated adequate internal consistency in both out-patient and non-clinical samples and adequate external validity against other well-validated measures of depression and anxiety (Barkham et al., 1996; McEvoy, Burgess, Page, Nathan, & Fursland, 2013). In this study, clients completed the IIP-32 prior to their intake assessment. Clients’ IIP-32 total scores are used in primary analyses in order to reflect overall self-reported interpersonal vulnerability. Examination of the internal consistency of the IIP-32 items yielded a standardized Cronbach’s α coefficient of .84, indicating good internal consistency across the individual items that comprise the total score to be used below.

**Personality inventory for DSM-5-brief form (PID-5-BF).** The PID-5-BF (Krueger, Derringer, Markon, Watson, & Skodol, 2013) is a 25-item self-report personality trait assessment scale for adults. It assesses five personality trait domains including negative affect, detachment, antagonism, disinhibition, and psychoticism, with each trait domain consisting of 5 items. The measure also yields a total score, with higher ratings signifying greater overall personality dysfunction. Each item asks the respondent to indicate how well the item describes him or her generally and is rated on a 4-point Likert-type scale ranging from 0 = very false or often false to 3 = very true or often true. The total score has been found to be reliable and clinically useful to clinicians in DSM-5 field trials (Krueger et al., 2013), and a few more recent independent studies have found that the five subscales (all αs > .70; Untalan, 2015) and the total score (αs > .83; Fossati, Somma, Borroni, Markon, & Krueger, 2015) demonstrated satisfactory internal consistency. In this study, clients’ completed the PID-5-BF prior to their intake
assessment. Clients’ PID-5-BF total scores are used in primary analyses in order to reflect overall self-reported personality dysfunction related to interpersonal vulnerability. Examination of the internal consistency of the PID-5-BF items yielded a standardized Cronbach’s \( \alpha \) coefficient of .83, indicating good internal consistency across the individual items that comprise the total score to be used below.

**Thin slice observer (TS-observer).** The TS-observer is a 20-item observer rated measure, adapted from Sasso and Strunk (2013) that is used to assess others’ quick first impressions of clients. As described in detail below, independent raters used the TS-observer to rate thin slice video-clips taken from clients’ intake therapy assessments. The TS-observer captures raters’ first impressions of a number of client traits including clients’ willingness to engage in therapy, communication skills, and personality traits. All items on the TS-observer are rated on a 7-point Likert type scale ranging from 1=“don’t agree at all” to 7=“completely agree.” The first three items ask raters to gauge a clients’ potential fit for CT, given the clients’ descriptions of themselves in TS video-clips (as detailed below in Thin Slice Video Clips; items adopted from Sasso & Strunk, 2013). The next five TS items assesses each of the Big Five personality traits (i.e., extraversion, agreeableness, conscientiousness, neuroticism, and openness; items adopted from Sasso & Strunk, 2013; Oltmanns et al., 2004). Next, two single items, also adopted from prior thin slice studies (see Friedman et al., 2007; Oltmanns et al., 2004) ask raters to assess each clients’ likeability and physical attractiveness respectively. The next six items were used to assess each of three PDs (i.e., two items each for Avoidant, Histrionic, and Schizoid), which have been found to be discernible to thin slice raters in previous studies.
(Friedman et al., 2007; Oltmanns et al., 2004; Sasso & Strunk, 2013). Sasso and Strunk (2013) used the above items to comprise the TS-observer measure and from this measure examined 10 theoretically derived TS-subscacles. While the items of the original TS-observer were designed to assess theoretically distinct constructs, in this study I examine the possibility that raters are actually assessing a smaller number of interpersonal constructs, by conducting an Exploratory Factor Analysis (EFA; see Results section) on all of the original TS-observer items above and four newly added items aimed at further capturing interpersonal traits. For primary analyses below, specific hypotheses are made regarding the factor analytically derived subscales that result from the EFA. Additionally, as part of this measure and the factor-analytically derived subscales identified below are newly developed for use in this study, I evaluate the reliability and validity of the factor analytically derived subscales in the results section.

**Thin slice therapist (TS-therapist).** The TS-therapist form was created for this study in order to capture therapists’ quick first impressions of their clients, prior to meeting them and based solely on thin slice video clips (detailed below). Prior to session one and subsequent to being assigned a given client, therapists watched their assigned client’s thin slice video clip and rated their quick first impressions of the client using the TS-therapist measure. The items contained in the TS-therapist measure are identical to those contained in the TS-observer measure detailed above, except the omission of the physical attractiveness item. In the results section, I conduct an EFA on all of the TS-therapist items in order ascertain the number of distinct latent constructs evident to

---

*Newly added items include three items to assess clients’ communication skills and one item to assess raters’ perceptions of how difficult the client will be to work with.*
therapists in thin slice video clips. For primary analyses below, hypotheses are made
regarding the factor analytically derived subscales that result from the EFA.

*Standardized Interaction Task-Observer (SIT-observer).* As detailed above, the
SIT contains 17 brief (≤ 90s) interpersonal role-plays, each with simple instructions
regarding an interpersonal situation the client and assessor are to engage in and an
interpersonal behavior that the client is expected to show. I adapted\(^{10}\) the SIT-observer
measure from the Leising and colleagues (2011) study in order to assess others’
perceptions of clients’ interpersonal behaviors and maximum competence in showing
such behaviors. In addition to simple instructions for the client regarding the
interpersonal behavior they are trying to show, each of the 17 SIT tasks also contains a
set of four standard answers, which assessors used in order to respond to clients in a
manner that standardized the situational circumstances as much as possible across clients.
Some standard answers were also designed to make the tasks more challenging. The SIT-
observer ratings were obtained by having independent raters watch these video-recorded
interpersonal role plays, between client and assessor, which were done at each client’s
intake assessment. A small minority of clients were unwilling to do particular SIT tasks,
reporting that they felt unable to and/or to uncomfortable completing such tasks; however
no client skipped all SIT tasks. In the results section below, I describe the skip rate for
each task across all clients and how this missingness was resolved analytically. For all
primary analyses below, SIT-observer ratings on all 17 tasks were averaged to reflect

\(^{10}\) The interactions which SIT-observers watched in order to provide their ratings were identical to those
done in Leising et al. (2011) with the exception of one minor adaptation: the standard assessor response
options for task #7 “Be self-critical” were limited in order to minimize clients’ emotional distress.
clients’ overall SIT performance. Inter-rater reliability of the SIT-scores is reported below in the results section.

**Thin Slice (TS) and SIT Video Clips: Content and Availability**

**TS clips.** At each client’s intake, the first activity they were asked to participate in was the thin slice task. The task was presented to the client as three brief questions that served both a research function and a means of getting acquainted with the assessor. Clients were given 90 seconds maximum to respond to each question. All clips were less than 5 minutes in length and contained clients responses to the following three prompts: (1) “Who have been the important people in your life.” (2) “How do you spend your free time?”, and; (3) “How would you describe yourself as a person?” Of the 126 clients who completed an intake and were enrolled in the sample, the TS clip was unavailable for one of them, as their intake was not recorded. Of the 125 clients with TS clips available for rating by observers and therapists, observers rated all 125 and therapists rated 123.

**SIT clips.** At each client’s intake assessment they were asked to participate in the SIT role plays. They were told that this exercise would take about a half an hour and involve them role playing a series of very brief interactions with the assessor. The tasks were presented as a means of better assessing each client’s interpersonal skills in various contexts. Clients were handed a booklet containing the instructions (not including the assessor’s standard responses) for each of 17 role plays. They were informed that each task described a hypothetical interpersonal situation they would be engaging in with the assessor and a behavior they would be trying to show. They were also told that each task would take 90 seconds at most and that some of the things the assessor would say may
seems a bit challenging or even a bit upsetting but that this was all part of the procedure. All tasks were video recorded, leaving all SIT clips ≤ 30 minutes in length (minimum length observed = 15 minutes; maximum length = 30 minutes). Of the 126 clients who completed an intake and were enrolled in the sample, the SIT clip was unavailable for six of them, as this last portion of their intake was not recorded (N = 4) or the recording available was damaged (N = 2). Observers rated the SIT clips for all 120 clients for whom they were available.

**TS and SIT Rating Procedures**

Training for the thin slice ratings was brief (2 hours) and designed to ensure that raters had a shared understanding of the scale anchors and the meaning of constituent items (e.g., a shared definition of extraversion). At this initial training, raters were also briefed in confidentiality and lab procedures. Training for the SIT ratings was a bit more extensive (8 hours), as training was designed to gain a shared understanding of what a "good" and a "bad" performance constituted for each of the 17 tasks. Raters watched mock SIT interactions and good and bad anchors were evaluated with reference to how a socially skilled individual might approach a given task. For example on the “Assert a Claim” task, raters were taught to look for signs of effective communication including respecting one’s own needs and values while balancing the need to respect others.

Following training, five undergraduate raters provided TS and SIT ratings of the selected evaluation clips (i.e., SIT and TS clips from intake evaluations). Raters were randomly assigned to code both the TS and SIT clip for a given client. Thin slice clips were always rated prior to SIT clips, to ensure that TS ratings indeed constituted raters’
first impression of the client. For all analyses, the average of all available ratings made for each client will be utilized. All raters were blind to the study rationale and hypotheses.
Chapter 3: Results

Overview

The results are organized to address, in order, the three main objectives and corresponding hypotheses detailed in the Purpose of This Study section. Before reporting on these analyses, I provide information on the reliability of the observer ratings used in this study as well as the psychometric properties of original and adapted measures (i.e., CC, TS-therapist, TS-observer, and SIT-observer). I then present descriptive statistics for the outcome and predictor variables (i.e., all interpersonal vulnerability variables and symptoms) to provide a context for understanding findings of interest.

Psychometric Properties of Original and Adapted Scales

TS-observer. As discussed above in the measures section, one preliminary aim of this study is to better understand the interpersonal constructs that TS raters are actually assessing when responding to brief TS video clips using the TS-observer measure. While the items of the TS-observer were designed to capture ten theoretically distinct traits, it is possible that TS-ratings really reflect a smaller number of constructs, discernable to thin slice raters. To assess this possibility I conducted an EFA on all 20 items contained in the adapted TS-observer measure used in this study.

To determine the number of factors, I examined the scree plot and conducted a parallel analysis. The scree plot suggested a one or three factor solution. The parallel
analysis (see Appendix B: Figure 1) suggested three factors. To further evaluate the appropriate number of factors, I ran an EFA with a 3 factor solution and an oblique quartimin rotation. An oblique quartimin rotation was chosen as it allows factors to be correlated. As shown in Table 2, the three factors\(^{11}\) mapped on to theoretically meaningful constructs. Ten variables\(^{12}\) loaded onto the first factor, which appeared to reflect pro-treatment behaviors that may convey more willingness, interest, and ability to engage in treatment. The seven\(^{13}\) variables loading on factor two assess clients’ social tendencies including extraversion, perceived social network, and one’s proclivity for seeking attention. Lastly, three\(^{14}\) items loaded on factor 3, which appeared to reflect overall neuroticism. The inter-factor correlation matrix showed that the correlations among factors ranged from -.23 to .35. For use in all subsequent analyses, clients’ average score across the items of each factor-analytically derived TS-observer subscale were calculated. Prior to doing so, items were re-keyed as necessary so that all items being averaged were keyed in the same direction. The resulting scores reflect: (1) TS-observer: Pro-treatment behaviors; (2) TS-observer: Sociability, and; (3) TS-observer: Neuroticism.

\(^{11}\) Constituent items comprising each factor were determined by both factor loadings and theoretical consistency. For the majority of items, this approach resulted in items loading on to the factor for which they had the strongest loading. Two exceptions to this were Conscientiousness and Openness, which were grouped in to the “Pro-treatment behaviors” and “Sociability” factors respectively, despite loading most strongly onto the “Neuroticism” factor.

\(^{12}\) Pro-treatment behavior items: Treatment interest, No treatment reservations, Won't dropout of treatment, Agreeable, Conscientious, Likeable, Uses nonverbals well, Speaks well when communicating, Good conversation skills, Will be difficult to work with.

\(^{13}\) Sociability items: Extraverted, Openness, Attractive, Prefers to be alone, No close friends, Unhappy when not center of attention, Uses appearance to draw attention

\(^{14}\) Neuroticism items: Neuroticism, Worries about criticism from others, and Believes they are inferior.
To assess the inter-rater reliability for each of the TS-observer subscales, I calculated intra-class correlation coefficients (ICCs) estimating the variance attributable to the random effects for clients and assessors divided by the total of these components and variance attributable raters and error. ICCs were corrected for the harmonic mean (M = 4.82) of the number of ratings available per client. The ICCs were as follows: .82 for TS-observer: Pro-treatment behaviors, .84 for TS-observer: Sociability, and .63 to TS-observer: Neuroticism. Standardized Cronbach’s alphas were .94 for Pro-treatment behaviors, .81 for Sociability, and .77 for Neuroticism, indicating good to excellent internal consistency.

**TS-therapist.** As detailed above, the TS-therapist form was created for this study in order to capture therapists’ quick first impressions of their clients, prior to meeting them. Items of the TS-therapist measure were identical to the TS-observer measure, except that the physical attractiveness item was omitted. Because it is possible that therapists were able to decipher more distinct constructs based upon TS video clips than observers, I conducted an EFA on all 19 of the TS-therapist items.

To determine the number of factors, I examined the scree plot and conducted a parallel analysis. The scree plot suggested a one or four factor solution. The parallel analysis (see Appendix B: Figure 2) suggested four factors. To further evaluate the appropriate number of factors, I ran an EFA with a 4 factor solution and an oblique quartimin rotation. An oblique quartimin rotation was chosen as it allows factors to be correlated. The inter-factor correlation matrix showed that correlations among factors ranged from .33 to -.14. The four resulting factors also mapped on to theoretically related
constructs. The first factor contained the same ten variables as the TS-observer: Pro-treatment behaviors subscale identified above. Four variables loaded on to factor two, which captures clients’ social tendencies. All four of these items were also identified as mapping on to the TS-observer: Sociability subscale; however, the two items assessing Histrionic PD traits (i.e., Unhappy when not the center of attention and Uses appearance to draw attention to oneself) that were included in this “Sociability” subscale when observer ratings were examined loaded strongly onto their own distinct factor (Factor 4: TS-therapist: Histrionic) when therapist ratings were examined. Lastly, the structure of Factor 3: Neuroticism, identified from the TS-observer EFA, replicated exactly with TS-therapist ratings. For use in all subsequent analyses, clients’ average score across the items of each factor-analytically derived TS-therapist subscale were calculated. Prior to doing so, items were re-keyed as necessary so that all items being averaged were keyed in the same direction. The resulting scores reflect: (1) TS-therapist: Pro-treatment behaviors; (2) TS-therapist: Sociability; (3) TS-therapist: Neuroticism, and; (4) TS-therapist: Histrionic. Examination of the internal consistency of the items of the four TS-therapist subscales yielded standardized Cronbach’s alpha coefficients of .89 for Pro-treatment behaviors, .68 for Sociability, .66 for Neuroticism, and .75 for Histrionic PD traits.

SIT ratings. As noted above, a small minority of clients were unwilling to do particular SIT tasks. The majority of SIT tasks were never skipped by any clients, and the vast majority (N = 94, 78%) of clients completed all 17 tasks. Only three tasks were
skipped by more than three clients, with these tasks including: “Express Affection”, “Ask for Emotional Support”, and “Present Yourself Positively”. A total of 26 clients skipped at least one task. Among these, 16 clients skipped just one task, 7 skipped two tasks, 1 skipped three tasks, 1 skipped four tasks, and 1 client skipped five tasks.

To address these missing values single imputation was used. The missing values were deemed to be missing at random (MAR), as the missingness of these SIT task scores was significantly related to other measured client variables, including clients’ scores on other SIT items they did complete, TS-scores, and Social Phobia diagnosis (for a more in-depth discussion of MAR data, see Steyerberg, 2008). Given this, I chose a maximum likelihood (ML) approach, specifically expectation-maximization or EM, in order to impute missing values for individual SIT items (Enders, 2001; Steyerberg, 2008). ML approaches utilize the mean vector and covariance matrix from available cases to create an imputed dataset. The way that ML estimates the parameters in this imputed dataset depends on the algorithm selected. The EM algorithm used here, first estimates the missing value of a given variable, given the observed values in the data, and then processes this data using ML estimation to give new mean and covariance estimates. EM then uses these new estimates to again re-estimate the missing values. This process is repeated until the maximum change in estimates does not exceed a certain criterion (Yim, 2015). ML methods, such as EM, have been shown to deal with higher missing percentages in a manner that yields less biased estimates as compared to other methods including list-wise deletions and stochastic imputation regression methods (Yim, 2015).

---

16 I used convergence criteria of .0001, meaning that the estimation procedure was continued until the change in parameter estimates between iterations was less than .0001 for each parameter (i.e., means and covariances). Convergence was achieved in ten iterations.
For all analyses below, the SIT-observer predictor reflects each client’s average score across all 17 tasks, including imputed values\(^{17}\) on any skipped tasks.

To assess the reliability of SIT observer ratings, I calculated intra-class correlations (ICCs) to estimate the variance attributable to the random effects for clients and assessors and divided the total of these components by the variance attributable to the total of these factors along with raters and error. ICCs were corrected for the harmonic mean (M = 4.82) of the number of ratings available per client. The ICC for the SIT-observer score was .89 indicating raters achieved good inter-rater reliability. Examination of the internal consistency of the 17 individual items averaged to create the SIT-observer score yielded a standardized Cronbach’s α coefficient of .95, indicating excellent internal consistency across the individual items. In the descriptive statistics section, the convergent validity of SIT-observer is discussed.

**In-session cognitive change (CC).** As detailed above, the CC measure was created for this study in order to capture the extent to which clients re-evaluated negative automatic thoughts or core beliefs within the session that just occurred (i.e., broad assessment of in-session cognitive change). To determine the appropriate scoring for the CC, I examined the scree plot and conducted a parallel analysis. The scree plot suggested a one factor solution, given the large drop off in eigenvalues after the first factor. The

\(^{17}\)The SIT-observer score varied minimally as a function of being calculated from constituent tasks with imputed values (M = 5.30, SD = .90, Range = 2.26-6.72) vs being calculated from constituent tasks without imputed values (M = 5.31, SD = .90, Range = 2.26-6.72); nevertheless, the SIT-observer score used below was calculated using imputed values on any skipped tasks. The main finding for the SIT-observer score depicted in Table 5 is nearly identical if the SIT-observer score, calculated with missing values on skipped SIT-tasks, is used instead (effect on slope of symptom change: \(b = -.16, SE = .07, t\)-value = -2.32, \(p = .02\).
parallel analysis (see Appendix B: Figure 3) suggested one factor as well. For all analyses below, clients’ scores on all five CC items were summed to reflect their total CC-score. Examination of the internal consistency of the individual items summed to create the CC-score yielded standardized Cronbach’s α coefficients ranging from .85 to .93, across each of the first five therapy sessions. In the descriptive statistics section, the convergent validity of the CC measure is discussed.

**Descriptive Statistics and Change in Depressive Symptoms**

**Change in depressive symptoms.** With regard to changes in the main outcome, depressive symptoms, clients experienced an average improvement of 4.47 (SD = 8.74) BDI-II points from session 1 to 5 ($t = 5.60, p < .0001, d = 1.03$) and an average improvement of 15.23 (SD = 12.12, $t = 13.99, p < .0001, d = 2.52$) BDI-II points from session 1 to their post-treatment assessment (or last recorded BDI-II score). Additionally, clients’ slope of symptom change across the course of treatment was significantly negative ($b = -1.07, SE = .06, t = -16.90, p < .0001$) indicating that clients, on average experienced significant symptom reductions across treatment with BDI-II scores decreasing by just over 1 point per week, on average.

**Descriptive statistics.** As a preliminary first step, I examined the relationships among all candidate predictors and baseline depressive symptoms (see Table 3). The majority of TS ratings, as assessed by both therapist and observer, did not show significant relations to depressive symptoms but were correlated in the expected direction (i.e., “positive” traits of sociability and pro-treatment behaviors were negatively correlated with depressive symptoms). TS: Neuroticism ratings were one exception to
this, with both observer and therapist ratings of neuroticism being significantly correlated with baseline depressive symptoms in the expected (i.e., positive) direction. SIT-observer ratings, although related in the expected negative direction, did not correlate significantly with depressive symptoms. Clients’ PID-5-BF and IIP-32 scores were significantly correlated with depressive symptoms, with clients higher in baseline depressive symptoms reporting significantly greater personality pathology and more interpersonal problems. With regard to relations among candidate predictors, TS ratings from observer and therapist informants on the three overlapping subscales were significantly correlated ($r$ ranging from .32-.62). TS-subscale ratings, across both informants, also showed expected relations to self-report measures and SIT-observer scores, with clients perceived to be higher in pro-treatment behaviors and sociability scoring more favorably on the SIT, PID-5-BF, and IIP-32. Clients’ PID-5-BF and IIP-32 scores were highly correlated ($r = .57$) indicating that clients’ self-reported personality pathology and interpersonal problems overlap considerably. While clients’ SIT scores showed expected negative relations with PID-5-BF ($r = -.15$) and IIP-32 ($-.09$) scores, they appeared to be somewhat independent of this self-report data as these relations were small and nonsignificant.

In order to assess the relationship between the key measures of interpersonal vulnerability examined in this study and common comorbid anxiety disorder diagnoses at baseline (i.e., social phobia and generalized anxiety disorder; GAD), I conducted a series of independent sample t-tests to detect any significant mean differences between those with and without these diagnoses on all key predictors. Results are depicted in Table 4.
Scores on thin-slice variables, as rated by both observers and therapists, did not differ significantly as a function of either social phobia or GAD diagnoses. SIT, PID-5-BF, and IIP-32 scores were significantly related to social phobia diagnosis, such that clients without such a diagnosis scored more favorably on these measures (i.e., higher SIT performance ratings, lower personality pathology, and fewer interpersonal problems on average). With regards to GAD diagnoses, only PID-5-BF scores differed significantly as a function of GAD comorbidity, such that those without GAD diagnoses had lower levels of personality pathology on average.

In order to assess the validity of the in-session cognitive change scale used in this study, I examined correlations between clients’ CC scores and scores on a previously studied measure of related CT skills, the Competencies of Cognitive Therapy Scale-Therapist Report (CCTS-TR; Strunk, Hollars, Adler, Goldstein, & Braun, 2014). I selected the CCTS-TR measure as it possess sound psychometric properties (see Strunk et al., 2014), including reliability and construct validity, with prior research showing that the CCTS-TR is significantly correlated with well-established measures (Ways of Responding Questionnaire; Barber & DeRubeis, 1992; DAS; Weissman & Beck, 1978) of CT skills and related cognitive constructs. The CCTS-TR is a brief measures completed by a client’s therapist that assesses client’s ability, frequency and independence of use in three different CT skill domains: automatic thoughts, behavioral activation, and core beliefs. In this study, both the CCTS-TR and the CC were administered following clients first treatment session. The correlation among the CCTS-TR total score and the CC total score at this time point was small, but statistically
significant \( r = .18, p = .04 \), indicating the clients who were perceived by therapists to be more skilled also reported experiencing more cognitive change during the first session. The CCTS-TR was also completed by therapists during clients’ fourth week of treatment, thus I also examined the correlation between clients’ CCTS-TR total score at week 4 and their average CC score across the first five therapy sessions and found a strong positive relationship \( r = .38, p < .0001 \).

**Objective 1: Interpersonal Vulnerability Predicting Treatment Outcome**

Objective one was to examine the possibility that observer rated interpersonal vulnerability will significantly predict both subsequent symptom change (Hypothesis 1A) and dropout (Hypothesis 1B) and that these effects will be numerically stronger than those observed for self-reported interpersonal vulnerability. Another initial objective was to examine the effects of TS-therapist ratings on symptom change (Hypothesis 2A) and dropout (Hypothesis 2B).

**Hypothesis 1A.** In order to examine the relationship between baseline interpersonal vulnerability and subsequent symptom change throughout treatment, I used a Hierarchical Linear Model (HLM) implemented in SAS PROC MIXED, to model person-specific intercept and slope parameters in BDI-II scores over time. The primary effects of interest in these models were the interactions between each interpersonal vulnerability predictor and time, with the interpersonal vulnerability predictors being: SIT-observer mean score, TS-observer: Pro-treatment behaviors, TS-observer: Sociability, TS-observer: Neuroticism, the IIP-32 total score, and the PID-5-BF total score. I predicted that the numerically strongest predictors of slope of symptom change in
this HLM model (Hypothesis 1A) would be the four observer ratings (i.e., SIT and TS scores).

Results are depicted in Table 5, along with an additional value: $b^*\star$, or the beta obtained for the predictor when all of the predictors were first standardized ($M = 0$, $SD = 1$) prior to being entered in the model. As the table shows, only PID-5-BF scores and SIT-observer scores significantly predicted the slope of symptom change. PID-5-BF self-report scores were the numerically strongest predictor of slope, with self-reported personality pathology predicting a steeper slope of symptom change throughout treatment (indicated by the negative $t$-value in the table). This runs counter to my hypothesis that observer rated interpersonal vulnerability measures would be the numerically strongest predictors of symptom change. Consistent with my hypothesis, clients with higher (i.e., more favorable) SIT-observer ratings had significantly steeper slopes of symptom change throughout treatment. None of the TS-observer rated constructs nor clients’ IIP-32 self-report scores showed significant relations to slope of symptom change.

**Hypothesis 1B.** In order to identify interpersonal predictors of dropout, I fit a logistic regression model examining the effects of the following: SIT-observer mean score, TS-observer: Pro-treatment behaviors, TS-observer: Sociability, TS-observer: Neuroticism, the IIP-32 total score, and the PID-5-BF total score. I predicted that the numerically strongest predictors of dropout (Hypothesis 1B) would be: the mean SIT-observer score and TS-observer: Neuroticism. Results are depicted in Table 6. Odds ratios from this model indicate the change in odds of dropout associated with a 1 SD unit change in the predictor variable. Counter to my predictions, neither the observer-rated nor
the self-reported measures of interpersonal vulnerability emerged as significant predictors of dropout (all $p$s > .38).

**Hypothesis 2A.** In order to examine the relation between TS-therapist ratings and subsequent slope of symptom change I used HLM to model person-specific intercept and slope parameters in BDI-II scores over time. The primary effects of interest in these models were the interactions between each of the following and time: TS-therapist: Pro-treatment behaviors, TS-therapist: Sociability, TS-therapist: Neuroticism and TS-therapist: Histrionic. I predicted that the TS-subscales identified as predictors when observer ratings were examined would also be significant predictors of symptom change when assessed by therapists. Given the null findings observed for the effects of TS-observer ratings (see Table 5) I had no other specific hypotheses for relations of TS-therapist ratings and outcome.

Results are depicted in Table 7. As depicted in the table, the only significant predictor of slope of symptom change was TS-therapist: Neuroticism ratings, such that those perceived by therapists to be higher in neuroticism pre-treatment tended to experience steeper slopes of symptom change throughout treatment$^{18}$ ($p = .02$). In addition to significantly predicting a steeper slope of symptom change, TS-therapist: Neuroticism ratings also showed a trend ($p = .06$) relation to clients’ BDI-II intercept such that those perceived by therapists to be high in neuroticism had higher depressive symptoms at their intake evaluations. Lastly, TS-therapist: Histrionic ratings were related to clients’ BDI-II intercepts, such that those perceived to be higher in Histrionic like traits

---

$^{18}$ This effect of TS-therapist: Neuroticism on slope of symptom change was still significant when self-report measures of interpersonal vulnerability (i.e., PID-5-BF, IIP-32) and the interactions between these variables and time were included as covariates.
tended to have significantly less severe depressive symptoms at their intake evaluation ($p = .04$).

**Hypothesis 2B.** In order to examine the relation between TS-therapist ratings and treatment dropout, I fit a logistic regression model examining the effects of all four factor-analytically identified TS-therapist subscales. I originally predicted that the TS-sub scales identified as predictors when observer ratings were examined would also be significant predictors of dropout when assessed by therapists. As shown in Table 8, TS-therapist predictors had non-significant effects on dropout, as was the case for TS-observer ratings.

**Exploratory analyses: Informant differences.** In exploratory analyses, I planned to test whether relations between TS subscales and slope of symptom change varied as a function of whether the TS rater was a therapist or an observer, by examining the effect of interactions between TS subscales of interest and “rater type” in predicting slope of symptom change. Insofar as therapists’ unique perspectives have important consequences for the process of change in CT, I would expect therapist ratings may predict more strongly than observer ratings. To address this question analytically, I created one stacked thin slice variable for each of the overlapping TS subscales (i.e., TS: Pro-treatment behaviors, TS: Sociability, and TS: Neuroticism). Specifically, for a given TS subscale this stacked variable contained both TS-therapist ratings and TS-observer ratings for each client. I created a corresponding “rater type” variable that appropriately flagged ratings in the stacked variable as being from therapist or observer. Then, for each TS construct I examined the effects of a three-way interaction term (TS-construct*rater
type*time) on symptom change over the course of treatment. Results are depicted in Table 9. The null findings indicate that the effects of the examined TS-constructs on clients’ intercept and slope of symptom change throughout treatment do not vary meaningfully as a function of whether the rater was a therapist or an observer.

**Objective 2: Interpersonal Vulnerability Predicting Key CT Processes**

Objective two was to examine the possibility that baseline interpersonal vulnerability, as assessed by TS- and SIT-observer ratings, has significant effects on the formation of the therapeutic alliance (Hypothesis 3) and clients’ in-session cognitive change (Hypothesis 4) early in treatment. Prior to examining these hypotheses, I assessed the relationship among clients’ CC and WAI scores separately at each of the first five therapy sessions. On average, WAI and CC scores were significantly correlated at these early time points (average \( r = .52 \))\(^{19}\).

**Hypothesis 3.** To examine the possibility that others’ perceptions of clients’ heightened interpersonal vulnerability at baseline significantly predicts lower average therapeutic alliance scores, I used HLM to model person-specific intercept and slope parameters in Alliance. The primary predictors of interest in these models were the following (both the main effects of each predictor and interactions between each predictor and time): SIT-observer, TS-observer: Pro-treatment behaviors, TS-observer: Sociability, and TS-observer: Neuroticism. Clients’ baseline depressive symptom severity was also controlled for. I predicted that heightened interpersonal vulnerability, as

---

\(^{19}\) To obtain this average correlation I examined the correlation between WAI and CC scores separately at each time point (all \( ps < .01 \)). I then converted these correlation coefficients to z-scores using Fisher’s r to z transformation, averaged these z-scores and converted this average back to an r-value.
assessed specifically by the TS- and SIT-observer ratings, would predict lower average therapeutic alliances, with exploratory analyses also examining relations between these scores and slope of alliance. Results are depicted in Table 10. Findings revealed partial support for my predictions, as one of the TS-observer ratings emerged as a significant predictor of average alliance scores. TS-observer ratings suggesting greater pro-treatment behaviors at baseline predicted higher average alliance scores across the early treatment sessions examined. There was no relationship between any observer-rated interpersonal variables and slope of alliance scores. Time was also a significant predictor, suggesting that clients’ alliance scores increased over this early phase of treatment.

In order to examine the possibility that the relation between TS-observer: Pro-treatment behaviors and average alliance scores identified was not better accounted for by potential confounds, I ran a follow-up model examining the effects of TS-observer: Pro-treatment behaviors on alliance while controlling for self-reported interpersonal vulnerability variables and concurrent symptom change (i.e., intercept and slope of BDI-II scores over the time points at which WAI was assessed). Results from this model are depicted in Table 11. Findings suggest that the effect of observer-rated pro-treatment behavior on average alliance scores is numerically stronger ($p = .007$) than the effects of self-reported interpersonal problems ($p = .10$) or self-reported personality pathology ($p = .19$). In addition to the significant effect of TS-observer: Pro-treatment behaviors on average alliance, results showed that average concurrent symptom severity was also a

---

20 In this model, and all subsequent models examining Objective 2, the time variable was person-mean centered so that the WAI and CC intercepts predicted for each client would reflect each client’s average WAI and CC score at the mid-point of the sessions examined.
strong predictor of average alliance scores, such that those with higher BDI-II scores tended to have lower therapeutic alliances.

**Hypothesis 4.** To examine the possibility that others’ perceptions of clients’ heightened interpersonal vulnerability at baseline significantly predicts lower in-session cognitive change scores early in treatment, I used HLM to model person-specific intercept and slope parameters in CC scores. The primary predictors of interest in these models were the following (both the main effects of each predictor and interactions between each predictor and time): SIT-observer, TS-observer: Pro-treatment behaviors, TS-observer: Sociability, and TS-observer: Neuroticism. Clients’ baseline depressive symptom severity was also controlled for. I predicted that heightened interpersonal vulnerability, as assessed specifically by the TS- and SIT-observer ratings, would predict lower average in-session cognitive change scores, with exploratory analyses also examining relations between these scores and slope of in-session cognitive change. Results are depicted in Table 12. Findings revealed partial support for my prediction, namely one of the TS-observer ratings emerged as a significant predictor of average in-session cognitive change. Specifically higher TS-observer pro-treatment behaviors at baseline predicted significantly higher average CC scores across the early treatment sessions. There was no relationship between any observer-rated interpersonal variables and slope of CC scores across these early sessions.

In order to examine the possibility that the relation between TS-observer: Pro-treatment behaviors and average in-session CC scores identified was not better accounted for by potential confounds, I ran a follow-up model examining the effects of TS-observer:
Pro-treatment behaviors on CC scores while controlling for self-reported interpersonal vulnerability variables and concurrent symptom change. Results from this model are depicted in Table 13. Findings suggest that the effect of observer-rated pro-treatment behavior on average in-session cognitive change is numerically stronger ($p = .001$) than the effects of self-reported interpersonal problems ($p = .12$) or personality pathology ($p = .03$). There was also a significant effect for self-reported personality pathology in predicting average CC scores, such that greater perceived personality pathology predicted greater cognitive change on average. Additionally, the significant main effect of time in predicting CC scores suggests that across these early sessions clients are experiencing an increasing amount of adaptive in-session cognitive changes.

**Exploratory analyses.**

**Predictive effects of therapist ratings on average alliance.** To explore the extent to which a therapist’s first impressions of clients’ interpersonal vulnerability predicts subsequent therapeutic alliance, I ran an HLM model examining the effects of the following (and their interactions with time) in predicting alliance: TS-therapist: Pro-treatment behaviors, TS-therapist: Sociability, TS-therapist: Neuroticism and TS-therapist: Histrionic. Clients’ baseline depressive severity was included as a covariate. Given my original prediction that those TS-constructs that showed significant relations to alliance when observer ratings were examined would also be the constructs that predicted when therapist ratings were examined, I expected that the strongest predictor of average alliance in this model would be TS-therapist: Pro-treatment behaviors. Results are depicted in Table 14. While the direction of the effect for TS-therapist: Pro-treatment
behaviors on average alliance was the same as that observed for the effect of TS-observer: Pro-treatment behaviors, the relationship failed to achieve statistical significance. TS-therapist: Sociability did predict average alliance, such that clients therapists perceived to be more sociable formed significantly greater alliances on average than those perceived to be less sociable. I ran a follow-up model examining the effects of TS-therapist: Sociability on alliance while controlling for self-reported interpersonal vulnerability variables and concurrent symptom change (see Table 15). Findings from this model suggest that the effect of therapist-rated sociability on average alliance scores is numerically stronger ($p = .002$) than the effects of self-reported interpersonal problems ($p = .45$) or self-reported personality pathology ($p = .42$). In addition to the significant effect of TS-therapist: Sociability, results showed that average concurrent symptom severity (i.e., average BDI-II scores over the sessions at which WAI was assessed) was also a strong predictor of average alliance scores, such that those with higher BDI-II scores tended to have lower therapeutic alliances.

**Predictive effects of therapist ratings on clients’ average in-session cognitive change.** To explore the extent to which a therapist’s first impressions of clients’ interpersonal vulnerability predicts clients’ subsequent cognitive change, I ran an HLM model examining the effects of the following (and their interactions with time) in predicting client-reported in-session cognitive change: TS-therapist: Pro-treatment behaviors, TS-therapist: Sociability, TS-therapist: Neuroticism and TS-therapist: Histrionic. Clients’ baseline depressive symptom severity was controlled for. Given my original prediction that those TS-constructs that showed significant relations to cognitive change...
change when observer ratings were examined would also be the constructs that predicted when therapist ratings were examined, I expected that the strongest predictor of average in-session cognitive change in this model would be the TS-therapist: Pro-treatment behavior variable. Results are depicted in Table 16. Contrary to my hypothesis, TS-therapist: Pro-treatment behavior did not significantly predict average in-session cognitive change ($p = .21$). No other TS-therapist predictors emerged as significant predictors of either average CC scores or slope of CC scores early in treatment. Time was the only significant predictor in the model ($b = .48, p = .0004$) indicating that in-session cognitive change was increasing over the early sessions.

**Variability in effects across TS informants.** In additional exploratory analyses, I examined how the above relations, between TS-ratings and therapeutic processes, varied across observer and therapist TS informants. To do so, I examined the effects of interactions between TS subscales of interest and “rater type” in predicting both alliance and in-session cognitive change (using separate models for each process measure). Results for these two HLM models are depicted in Table 17 (predicting alliance) and Table 18 (predicting in-session cognitive change). As can be seen from both tables, the null findings for interactions between TS-constructs and rater type indicated none of the relations of interest varied significantly as a function of whether the TS rater was a therapist or an observer (viz., the relation of TS constructs and average alliance, TS constructs and average in-session cognitive change, TS-constructs and slope of alliance, or TS-constructs and slope of cognitive).
Objective 3: Alliance and In-session Cognitive Change as Mediators

The final objective was to examine the possibility that clients’ therapeutic alliance and self-reported in-session cognitive change early in treatment might account for any significant relations between interpersonal vulnerability at intake and slope of symptom change. For the alliance mediation models, I planned to limit the mediational analyses to interpersonal constructs that showed significant relations to slope of symptom change and to the alliance. Similarly, for cognitive change mediation models, I planned to limit mediational analyses to interpersonal constructs that significantly predicted slope of symptom change and in-session cognitive change. The only interpersonal predictor that met these criteria, given the findings detailed above, was self-reported personality pathology as assessed by the PID-5-BF. PID-5-BF scores significantly predicted clients’ subsequent reductions (i.e., slopes) in depressive symptoms (b* = -.20, p = .007) and clients’ average self-reported in-session cognitive change (b* = 1.15, p = .04). PID-5-BF only met these criteria of significance for the cognitive change mediation model.

In order to test for a pattern of effects consistent with mediation involving PID-5-BF scores, CC scores, and outcome, I used the approach suggested by Preacher and colleagues (2010) for multilevel structural equation modeling (MSEM; Preacher, Zyphur, & Zhang, 2010). This MSEM approach overcomes limitations of traditional methods (Baron & Kenny, 1986) of assessing mediation which assume independence of observations, an assumption not upheld in repeated measures data in which within-client observations are correlated. Ignoring such clustering using ordinary regression can lead to downwardly biased standard errors. The mediational relationship I am testing can be
operationalized as a 2-2-1 model (where personality pathology at baseline and average early cognitive change across the first five sessions are at level 2 / the client level and depressive symptoms are at level 1 / nested within-clients). The MSEM approach also improves upon more recent multilevel modeling (MLM) methods (see Krull & MacKinnon, 2001; Pituch, Stapleton, & Kang, 2006) for testing mediation with 2-level nested data that involve non-simultaneous estimation using a two-step process (i.e., using Ordinary Least Squares regression to obtain path “a” of the indirect effect and MLM to obtain path “b” of the indirect effect). The non-simultaneous estimation of direct, indirect, and total effects used in these two step approaches can lead to biased estimates as the multiple components of variance are not considered simultaneously. In contrast, the simultaneous estimation of the parameters that make up the indirect effects used in the present MSEM approach allows for the consideration of multiple components of variance at once, reducing bias in the indirect effect estimate and allowing one to flexibly fit a variety of nested data structures. Following the specifications21 detailed by Preacher et al. (2010) I fit a 2-2-1 MSEM model examining the relationship between clients’ personality pathology and slope of symptom change as mediated by clients average in-

---

21 The model was specified as follows: X = PID-5-BF, M = CC scale and Y = BDI-II scores. X, M and Y were defined as latent variables, where indicators of M were clients’ average CC score, across all five sessions, on each of the individual items that comprise the CC total score (all loadings were fixed to 1 such that M would reflect clients’ average CC score across early sessions). X was defined by clients’ scores on each of the 25 items that comprise the PID total score with all loadings fixed to 1, such that X would reflect clients’ average PID score. Lastly, two latent variables were specified for Y, one reflecting BDI intercept another reflecting the BDI growth curve (i.e., similar to clients’ slopes). The intercept was defined by clients’ BDI total scores at intake and at each therapy session (all loadings fixed to 1 such that the intercept would reflect average symptom severity across all time points), and the slope latent construct was defined by clients’ total scores at intake and at each therapy session with all loadings fixed sequentially as specified in latent growth curve modeling approaches and per Dr. Preachers’ recommendation via personal communication (starting at 0 and increasing by 1 to the last session).
session cognitive change. The model was fit using the ‘growth’ function in the Lavaan Package in R.

Fit was poor for the cognitive change MSEM model as compared to established cutoffs (Hu & Bentler, 1999), with a root mean squared error of approximation (RMSEA) = .15 and a comparative fit index (CFI) = .47, suggesting that results from the model should be interpreted with caution. The estimate of the indirect effect of personality pathology on slope symptom change via in-session cognitive change was not significant ($b = -.05, p = .49, SE = .07$), suggesting that cognitive change does not mediate the relation between personality pathology and slope of symptom change. The “a” path (i.e., the effect of PID-5-BF on average cognitive change; $b = .27, p = .49, SE = .39$) was not significant indicating that PID-5-BF scores did not relate to average cognitive change (whereas PID-5-BF was significantly related to higher average cognitive change in the HLM analyses above, see Table 13. Thus, while the direction of the effect was consistent, the magnitude was not). The “b” path suggested a significant relationship such that higher average cognitive change was related to steeper reductions in symptoms ($b = -.17, p = .004, SE = .06$). Lastly, the “c” path (i.e., the effect of PID-5-BF on slope of symptom change or latent growth in BDI scores: $b = -.47, p = .009, SE = .18$) was consistent in magnitude and direction to the related effect identified in the HLM model detailed in Table 5.

---

22While the majority of clients’ BDI-II scores appeared to decrease linearly over the course of treatment, a quadratic latent growth curve was also modeled in an attempt to improve fit. Fit improved minimally in this model (CFI = .52; RMSEA = .15) and results remained unchanged from those reported in text.

23Given the poor fit of the MSEM model, I also examined this mediational hypothesis using the two-step approach detailed by Krull and MacKinnon (2001) and found no evidence of personality pathology effecting symptom change indirectly via in-session cognitive change (indirect effect: $b = -.0002, p > .05$).
In order to see how these relations varied when only symptom change occurring subsequent to the measurement of in-session cognitive change (i.e., session 6 onward) was modeled, I fit an additional MSEM model with the same specifications as the above, but only with BDI-II scores from subsequent time points included as indicators. The “b” path was no longer significant ($b = -.03, p = .72$); however, results for the indirect effect ($b = -.008, p = .74$), “a” path ($b = .27, p = .49$), and “c” path ($b = -1.07, p < .0001$) were similar in magnitude and direction to the effects found above.
Chapter 4: Discussion

In this study I examined the role of interpersonal vulnerability in CT for depression, with a specific focus on assessing these vulnerabilities from multiple perspectives. I found partial support for my hypothesis that clients’ interpersonal vulnerabilities would predict their symptom trajectories over the course of CT. One of the two client self-report assessments of interpersonal vulnerability, the PID-5-BF, emerged as a significant predictor of subsequent slope of symptom change, with clients who reported higher personality pathology at baseline experiencing more rapid symptom reductions across treatment. Additionally, client neuroticism, as perceived by therapists, and more favorable client role play performances, as perceived by observers, both predicted a steeper slope of symptom change. While the effect of SIT-observer ratings was numerically stronger than the effect of self-reported interpersonal problems, it was not numerically stronger than the effect of self-reported personality pathology. Both appear to be important, independent predictors of symptom change. With regard to client dropout, interpersonal vulnerability did not have any significant effects, contrary to my prediction that select observer and therapist rated constructs would be predictive. I found some evidence for my hypothesis that interpersonal vulnerability as assessed via SIT- and TS-observer ratings would predict lower average alliance and lower average in-session cognitive change, as clients perceived by observers to be lower in pro-treatment
behaviors tended to form a lower average alliance and experience less in-session cognitive change, with TS-observer: Pro-treatment behaviors showing incremental predictive validity over self-report measures of interpersonal vulnerability. Self-reported personality pathology, as assessed via the PID-5-BF, also demonstrated a significant relation to cognitive change, with clients higher in personality pathology experiencing significantly greater in-session cognitive changes. Lastly, I failed to find evidence to support the possibility that cognitive change or the alliance serve as mediators of identified relations between interpersonal vulnerability and outcome. An in depth discussion of the above findings follows below, but first, I briefly review the analyses conducted to examine the psychometric properties of the original and newly adapted scales used in this study and discuss relations among these constructs.

**Key Predictors: Psychometric Properties and Relations among Constructs**

**Distinct constructs discernable to thin slice raters.** Although not included as a formal hypothesis, one preliminary question of interest in this study was, what exactly do thin slice ratings reflect? While no investigation, to my knowledge, has claimed that thin-slice ratings of personality traits and interpersonal difficulties can serve as substitutes for diagnostic information, some assert that thin slice raters can accurately discern pathological personality traits (Lilienfeld & Patrick, 2009) associated with certain clinical diagnoses; whereas, others support the notion that such ratings are reliable proxies for global impressions of the target individual (Oltmanns et al., 2004; Sasso & Strunk, 2013) that are related to meaningful outcomes including those in the clinical domain. However, proponents of the latter position still assess these impressions using scales which yield
multiple “distinct” scores, thought to reflect several different interpersonal and personality traits. I conducted an EFA on both the TS-observer and TS-therapist measures, which are comprised mostly of items used in prior studies (see Oltmanns et al., 2004; Sasso & Strunk, 2013) in order to better assess the distinct latent constructs evident to thin-slice raters. An examination of the scree plots and the parallel analyses supported a three factor solution for the TS-observer scale and a four factor solution for the TS-therapist scale, suggesting that raters are able to discern a small number of distinct constructs. The same items loaded strongly onto the first factor (i.e., Pro-treatment behaviors) for both therapists and observers. Neuroticism and sociability were also discernable to both therapist and observers. Lastly, the number of factors identified across analyses suggests that therapists may be able to detect slightly more constructs relevant to personality pathology on the basis of thin slice clips, with histrionic traits being distinct from sociability (i.e., extraversion) like traits for therapists but not observers.

**Construct validity and relations among predictors.** Relations among all candidate predictors, baseline diagnostic info, and depressive symptoms (see Tables 3 and 4) provide some insight into the validity of the interpersonal vulnerability measures used. While all TS and SIT scales were associated with depressive symptoms in the expected directions, correlations were weak with only two (i.e., depressive symptoms and TS: Neuroticism assessed via therapist and observer) achieving significance, suggesting that TS and SIT scores are not merely proxies for clients’ symptom severity. Additionally, significant correlations among informants on overlapping TS-subscases suggest that therapists and observers quick first impressions tend to correspond with one
another. Self-reported interpersonal vulnerability and personality pathology, as assessed via the PID-5-BF and IIP-32, showed significant relations with depressive symptoms, replicating prior findings (Krueger et al., 2012; McEvoy, Burgess, Page, et al., 2013). With regard to anxiety disorder diagnoses, scores on thin slice predictors from both informants did not differ significantly as a function of social phobia or GAD whereas SIT, IIP-32, and PID-5-BF scores did (in anticipated directions, with mean SIT and IIP-32 scores only differing significantly as a function of social phobia). Taken together, the above findings suggest that observer and therapist thin slice ratings tend to have less overlap with clinical diagnoses and symptoms than self-report measures of related constructs. While shared method variance among client self-report measures offers one plausible explanation for this pattern of effects, it is also possible that observer and therapist TS ratings show weaker relations to depressive symptoms because they reflect distinct, yet clinically meaningful variability in client tendencies. This latter perspective highlights the potential for TS-ratings to serve as important compliments to traditional diagnostic and self-report information.

The CC form was created for the naturalistic treatment study examined here in order to provide a broad, easy to administer, assessment of in-session cognitive change. Preliminary findings supported the convergent validity of the CC scale with another established measure of clients’ cognitive skill use, the CCTS-TR (Strunk et al., 2014), suggesting that this new CC measure may provide a quick and valid means of assessing clients’ in-session cognitive change. However, it is important to note that while the relation between clients’ average CC scores across early therapy sessions and their
subsequent week 4 CCTS-TR score was strong, the initial correlation among these measures at session one was significant but relatively small.

Main Findings and Support for Hypotheses

Objective 1: Interpersonal vulnerability predicting slope of symptom change.

I found some evidence that TS and SIT ratings of interpersonal vulnerability significantly predict subsequent slope of symptom change, with both higher SIT-observer ratings and higher TS-therapist: Neuroticism ratings predicting more rapid decreases in depressive symptoms over the course of treatment. The relation between TS-therapist: Neuroticism and more rapid symptom change may seem at odds with findings showing that neuroticism tends to be related to heightened depressive symptoms (Kotov, Gamez, Schmidt, & Watson, 2010) and lower odds of treatment response (Quilty et al., 2008). However, treatment response is often determined as it was in the study by Quilty and colleagues (2008); by a combination of clients’ having experienced at least a 50% decrease in depressive symptoms pre-to-post treatment and clients’ final score on depressive symptom measures being lower than a pre-defined cutoff indicative of mild depressive symptoms. Thus, it is possible that clients higher in neuroticism may experience more rapid symptom reductions across treatment given they typically have more severe symptoms at baseline (i.e., regression to the mean) yet these same clients may meet responder status less frequently than those lower in neuroticism.

While I found that TS-therapist ratings of neuroticism predicted subsequent symptom change, TS-observer: Neuroticism ratings did not. One possible explanation for this difference in findings across informants is that there may have been more signal in
therapist as opposed to observer ratings of neuroticism, as therapists tended to utilize the full range of the scale more readily on the items composing this subscale (TS-therapist: Neuroticism subscale scores ranged from 1.33-7.00 with a SD = 1.13 whereas TS-observer: Neuroticism ratings ranged from 2.17-6.67 with a SD = .87).

The significant effect of SIT-ratings on subsequent slope of symptom change supports the notion that interpersonal behavioral deficits may be readily detected by observers and related to adverse outcomes in CT. The incremental predictive validity of SIT-ratings over and above self-reported interpersonal problems (i.e., IIP-32 scores) suggests that these behavioral tendencies, and the meaningful variability they predict, may not be fully captured via traditional self-report methods of assessing interpersonal problems.

Self-reported personality pathology (i.e., PID-5-BF scores) significantly predicted subsequent symptom change, with clients reporting greater PD pathology experiencing more rapid reductions in depressive symptoms over the course of treatment. Similar to the findings for TS-therapist: Neuroticism, this finding is seemingly inconsistent with the literature showing that those with PD pathology tend to have lower response and recovery rates in depression treatment and greater baseline depressive symptom severity (see Newton-Howes et al., 2006) than those with less or no PD pathology; however, this discrepancy may be explained in part by differing evaluations of both depression outcome and PD pathology. Findings relating PD pathology to outcome have been found to differ as a function of how personality pathology is measured and how depressive characteristics (e.g., chronicity, severity) are controlled for (Mulder, 2002). In this study
outcome was operationalized as slope of symptoms across the course of treatment, thus it is possible that clients in this study with higher PD pathology did experience more rapid reductions in symptoms, yet at the end of treatment did not have significantly lower symptom severity than those with lower PD pathology. This is also consistent with findings showing that those with more severe depressive symptoms tend to experience more rapid change and greater posttreatment effect sizes vs. those with less severe depression (see review by Driessen, Cuijpers, Hollon, & Dekker, 2010). It is important to note that overall I did not find full support for my hypothesis that observer and therapist ratings of interpersonal vulnerability would be numerically stronger predictors of subsequent slope of symptom change than related self-report measures, as PID-5-BF scores were the numerically strongest predictors of subsequent symptom change. This is somewhat at odds with the findings from Sasso and Strunk (2013) which highlighted the incremental validity of TS-observer ratings over and above self-report and diagnostic PD information. While PD pathology was assessed differently in the present study, the failure to replicate this incremental effect raises the possibility that TS-ratings may be less robust indicators of subsequent response than the initial investigation by Sasso and Strunk suggested.

Another main hypothesis regarding the relation among interpersonal vulnerability and treatment outcome was that observer rated interpersonal vulnerability would significantly predict dropout and that these effects would be numerically stronger than that of self-reported interpersonal vulnerability. Results revealed that neither self-report

---

24 I conducted a follow-up analysis examining PID-5-BF as a predictor of clients’ last observed BDI-II score and found no relation between PD pathology and final symptom severity ($b = .11, p = .36$)
nor observer rated constructs predicted dropout, a finding at odds with Sasso and Strunk (2013), who found that TS-observer: Neuroticism significantly predicted lower odds of dropout. The dropout rate in the sample investigated by Sasso and Strunk was relatively high compared to the dropout rate in the present sample (i.e., 33% vs. 17% in the present study), thus it is possible that the dropout analyses in the present study were underpowered in comparison (see Limitations section for a more in depth discussion). It is also possible that the relation of TS: Neuroticism and dropout identified by Sasso and Strunk is relatively weak or unreliable.

**Objective 2: Relations between interpersonal vulnerability and therapeutic processes.** With regard to predictors of the alliance, those perceived by TS-observers to be higher in pro-treatment behaviors and those perceived by TS-therapists to be higher in sociability formed significantly stronger alliances with their therapists, with both of these relations remaining significant when self-reported interpersonal vulnerability and concurrent symptom change were included as covariates. Additionally, the effects of both TS-observer: Pro-treatment behaviors and TS-therapist: Sociability ratings were numerically stronger than the effects of self-reported IIP-32 or PID-5-BF scores. Relations between TS-ratings and the alliance support the notion that quick first impressions predict the subsequent therapeutic relationship. The direction of these relations, such that more favorable impressions predicted better alliances, fits with prior findings of relations between maladaptive interpersonal styles, problematic attachment histories, and weaker alliances (Diener & Monroe, 2011; Holtforth et al., 2014). Interestingly the analyses in this study failed to find an association between SIT-observer
ratings and the alliance. To the extent that thin slice ratings reflect more global quick impressions or judgments whereas SIT-ratings reflect a more prescribed assessment of various maladaptive interpersonal behaviors, the pattern of effects observed here may suggest that the therapeutic relationship is more vulnerable to therapists’ internal judgments of clients than to therapists’ perceptions of clients’ readily observed behavioral deficits in the interpersonal domain.

With regard to predictors of in-session cognitive change, those perceived by TS-observers to be higher in pro-treatment behaviors reported significantly higher average CC scores across the early treatment sessions examined, with this effect remaining significant when self-reported interpersonal vulnerability and concurrent symptom change were included as covariates. While in the same direction as the findings for TS-observer ratings, none of the relations between TS-therapist ratings and in-session cognitive change reached significance, suggesting that the effect of TS: Pro-treatment behaviors on cognitive change may be unique to observers or a relatively weak effect. Future studies are needed to replicate this finding. Interestingly, there was also a significant effect for self-reported personality pathology on average CC scores, such that greater personality pathology predicted greater cognitive change on average, a finding somewhat counter to what one might expect given the rigidity of beliefs associated with PD pathology. One possible interpretation for this effect is that the quantity and rigidity of maladaptive beliefs in those with higher PD pathology also render them the most likely to experience cognitive change early in CT, as therapists have a wider variety of problematic schemas and thought patterns to target. Future studies are needed to examine
the extent to which this association between self-reported PD pathology and early cognitive change replicates and endures throughout the remainder of treatment.

**Objective 3: Therapeutic processes as mediators.** Overall, I failed to find support for my hypothesis that interpersonal vulnerability predicts subsequent symptom change indirectly via the therapeutic alliance or in-session cognitive change. Firstly, my primary analyses in objectives one and two identified only one (i.e., the PID-5-BF) interpersonal predictor that was significantly related to both subsequent symptom change and in-session cognitive change, suggesting that none of the observer nor therapist rated interpersonal predictors even met the basic “a”, “c” criteria required for mediation laid out by Baron and Kenny (1986) and others (Kraemer et al., 2002). An MSEM mediation model was implemented to examine in-session cognitive change as a mediator of the significant relation between PID-5-BF and subsequent slope of symptom change. The fit of the MSEM model was relatively poor, suggesting a few possibilities. One possibility is that my mediational hypotheses regarding how these constructs affect one another may not reflect the underlying structure of the data. Another factor that may have effected model fit is the sample size, with developers of the MSEM mediation approach recommending a minimum of 40 units or observations (per cluster-level indicator) to detect large (Cohen’s $d > .50$) structural paths at the between level and a minimum 100 units for smaller effects ($0.10 \leq d < .25$; Preacher, Zyphur, & Zhang, 2010). The majority of indicators in my analyses contained observations for roughly 120-126 units (i.e., clients); however, four of the BDI-II indicators at time points occurring late in treatment
contained closer to 100 observations (i.e., 99-105). Thus, I may have been slightly underpowered to detect a smaller mediational effect if it existed.

Although the fit was relatively poor for the MSEM model, the absence of a mediational effect of cognitive change was consistent with what I observed in the simpler two-step mediation model I conducted. These findings now add to the relatively limited literature examining cognitive change as a mediator of the relation between interpersonal vulnerability and outcome within CT for depression. Only one study has examined reductions in clients’ negative cognitions as a mediator of the relation between interpersonal vulnerability and outcome (McEvoy, Burgess, & Nathan, 2013) and found some evidence that concurrent pre-to-post treatment reductions in negative cognitions accounted for the relation between IIP scores and pre-to-post treatment BDI-II change. However, it is important to keep in mind that the analyses in the present study differ from those conducted by McEvoy and colleagues in terms of the operationalization of symptom change, cognitive change, analytic approach used, and treatment type. My failure to find evidence for the basic “a” and “c” criteria to support the alliance as a mediator is somewhat consistent with the mixed nature of prior findings, with some (McEvoy, Burgess, & Nathan, 2013) finding no evidence of interpersonal vulnerability predicting outcomes indirectly via the alliance and others finding some evidence of alliance mediating relations between interpersonal vulnerability at baseline and pre-to-post treatment symptom change (Hardy et al., 2001; Howard et al., 2006). Taken together, prior results coupled with the present findings suggest that cognitive change and the alliance may have rather small, or negligible, indirect effects on the association
between interpersonal vulnerability and slope of symptom change. Future work is needed to disentangle the strength and nature of these relationships, including the window of time in treatment during which they occur and the role of concurrent and prior symptom change.

**Limitations**

While this study highlights the value of using a multi-method approach to assessing interpersonal vulnerability in CT for depression, it is important to note several limitations. First, given the naturalistic observational design of this study, I cannot establish definitive causal relationships. Initial symptom severity is one important between-person difference that was modeled and/or controlled for in both the HLM and logistic regression models, suggesting that the effects observed were not better accounted for by depressive symptom severity. Additionally, interpersonal variables always temporally preceded the outcome they were predicting; however, I cannot rule out the possibility that the relationships observed between interpersonal predictors and subsequent outcomes were due to a separate, unmeasured variable. Additionally, had I found a significant indirect pathway in the MSEM models implemented, I would have only been able to conclude that the indirect effect of personality pathology on subsequent symptom change via therapeutic processes was consistent with a mediational relationship, as neither the alliance nor in-session cognitive change was experimentally manipulated and thus could not causally explain the relationship between PID-5-BF scores and outcome.
Second, my analyses examining relations between interpersonal vulnerability and CT processes focused on a window of sessions occurring relatively early in treatment and may not generalize to later therapy sessions. The decision to focus these analyses on early therapy sessions is supported by empirical findings showing that early response to CT (i.e., during the first four weeks) predicts treatment effectiveness (Tang & DeRubeis, 1999; Thase et al., 2000) and is estimated to account for roughly 60% of the total symptom reduction reported by clients (Busch, Kanter, Landes, & Kohlenberg, 2006; Fennell & Teasdale, 1987; Rush, Kovacs, Beck, Weissenburger, & Hollon, 1981). That being said, future studies are needed to examine the possibility that the effect of baseline interpersonal vulnerability on cognitive change and the therapeutic relationship may be more delayed, with adverse effects of clients’ vulnerabilities on these processes only becoming evident later in treatment. It is also possible that the effects of others’ perceptions of clients’ interpersonal vulnerability on both processes and outcome may be more immediate. Future studies assessing others’ perceptions of clients’ interpersonal vulnerabilities and key processes repeatedly over the course of treatment are needed to assess this possibility. Additionally, the findings in this study may not generalize to other forms of treatment for depression, as relations between personality pathology, related interpersonal vulnerabilities, and outcome have been shown to vary across different treatment modalities (Joyce et al., 2007; McBride, Atkinson, Quilty, & Michael, 2006).

Third, as discussed above, I may have been underpowered to detect mediational effects using the MSEM models implemented; however, the lack of evidence found for the mediational relations posited using either the MSEM or simpler two-step modeling
approach suggests that a mediational relationship among these constructs may not exist or be relatively small. Additionally, I may also have been underpowered to predict dropout. Indeed, some have suggested that a minimum of 10 “events” per parameter be present in logistic regression models in order to avoid problems of overestimated and underestimated variances and biased estimates of confidence intervals and coefficients; whereas others suggest that this “rule of 10” may be too conservative, finding a 5-9 events per parameter to be generally acceptable in more extensive simulation studies (see review by Hosmer, Lemeshow, & Sturdivant, 2013). Given the dropout rate in the present sample, the “rule of 10” would suggest including only two predictors in my logistic regression models. I conducted follow-up logistic analyses with just two predictors per model (i.e., BDI at intake and each predictor of interest), and found that neither self-report nor observer rated interpersonal vulnerability significantly predicted dropout (all \( ps > .48 \)).

Lastly, it is important to take into consideration the nature of thin slice and SIT ratings when interpreting these findings. While not significantly associated with baseline depressive symptoms, average SIT-ratings showed moderate to significant associations with clients’ social phobia diagnoses and self-reported personality pathology in the expected directions, yet were not significantly related to clients’ self-reported interpersonal problems. Consistent with findings from prior thin slice studies (see Oltmanns et al., 2004; Sasso & Strunk, 2013) correlations between thin slice ratings, self-reported personality pathology, and interpersonal problems were all in expected directions, but none were particularly strong. Several TS-therapist-self-report correlations
achieved significance but were relatively small in strength and only one TS-observer-
self-report correlation was significant. Thus, while observer-ratings of interpersonal
vulnerability are by no means perfect gauges of one’s self-reported interpersonal
difficulties or substitutes for traditional assessment measures, they do seem to provide
related yet distinguishable information that predicts meaningful variability in clients’
therapeutic experience. Taken together the findings from this study highlight the utility of
taking a multi-modal approach to assessing interpersonal vulnerability, as others’
perceptions of such traits provide an important complement to traditional assessment
methods.

**Conclusion**

These findings support the notion that individuals become aware of others’
interpersonal deficits and related personality traits on the basis of quick first impressions,
some of which have predictive validity in the clinical realm. Both thin slice ratings,
observer-ratings of role play performance, and self-reported personality pathology inform
clients’ likely rate of symptom reduction, pre-treatment. Additionally, thin slice ratings
and self-reported personality pathology provide valuable pre-treatment insight on the
likelihood of clients’ experiencing a weaker therapeutic alliance and less marked in-
session cognitive change. These findings regarding thin-slice ratings build on prior work
by Sasso and Strunk (2013) highlighting the predictive validity of quick first impressions
of personality traits and interpersonal tendencies in the clinical realm, and support the
potential of utilizing therapists’ quick first impressions. If these results are replicated,
therapists’ could potentially use thin-slice techniques in order to alert them to an
increased potential of suboptimal levels of symptom reduction with a given client, before treatment even starts. Indeed, several studies by Lambert and colleagues have found that providing feedback to therapists regarding clients’ clinical progress leads to improved outcomes and reduces client deterioration, particularly among those predicted to be non-responders (Lambert & Coco, 2014; Lambert, Harmon, Slade, Whipple, & Hawkins, 2005; Harmon et al., 2007). Taken together, a multi-modal approach to assessing interpersonal vulnerability pre-treatment may help clinicians to better identify at-risk clients.
References


Costa, P. T., & McCrae, R. R. (1992). *Revised NEO Personality Inventory (NEO-PIR) and NEO Five Factor Inventory (NEO-FFI) professional manual*. Odessa, FL: Psychological Assessment Resources


97


98


101


Appendix A: Tables
<table>
<thead>
<tr>
<th>Diagnosis</th>
<th>% of sample and number of clients (N) meeting criteria</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 MDE</td>
<td>44%(55)</td>
</tr>
<tr>
<td>≥2 MDE (Recurrent)</td>
<td>56%(71)</td>
</tr>
<tr>
<td>Any Axis I Comorbidity</td>
<td>71%(89)</td>
</tr>
<tr>
<td>GAD</td>
<td>25%(32)</td>
</tr>
<tr>
<td>Social Phobia</td>
<td>43%(54)</td>
</tr>
</tbody>
</table>

*Table 1. Primary and Comorbid Diagnoses and Specifiers at Intake.*

Note. MDE = Major Depressive Episode; GAD = Generalized Anxiety Disorder. Percentage of clients meeting criteria for each diagnosis is calculated out of the total sample at intake (n = 126). Comorbid diagnoses were not mutually exclusive categories, so some clients are represented in multiple categories.
Factor 1: Pro-treatment behaviors  | Factor 2: Sociability  | Factor 3: Neuroticism  
--- | --- | ---  
Treatment interest  | 91  | -13  | 19  
No treatment reservations  | 93  | -10  | 13  
Won't dropout of treatment  | 92  | -17  | 9  
Extraverted  | 42  | 45  | -12  
Agreeable  | 57  | 19  | -25  
Conscientious  | 29  | 21  | -33  
Neuroticism  | 16  | 13  | 62  
Openness  | 29  | 36  | -38  
Likeable  | 56  | 34  | -29  
Attractive  | -6  | 52  | -10  
Prefers to be alone  | -35  | -51  | 5  
No close friends  | -36  | -40  | 30  
Unhappy when not center of attention  | -1  | 69  | 33  
Uses appearance to draw attention  | -13  | 83  | 20  
Worries about criticism  | -2  | 16  | 70  
Believes they are inferior  | 0  | -7  | 81  
Uses nonverbals well  | 82  | 3  | 5  
Speaks well when communicating  | 81  | 9  | 3  
Good conversation skills  | 73  | 17  | -6  
Will be difficult to work with  | -79  | 8  | 26  

*Table 2. Rotated Factor Pattern from Exploratory Factor Analysis for Thin Slice-Observer Measure*

Note. Printed values are multiplied by 100 and rounded to the nearest integer. Loadings for variables mapping on to the given factor indicated in the column heading are bolded.
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>BDI</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>O: Pro-trx</td>
<td>-0.14</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>O: Social</td>
<td>-0.15†</td>
<td>.62**</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>O: Neurotic</td>
<td>.21*</td>
<td>-.17†</td>
<td>-.12</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>T: Pro-trx</td>
<td>-.09</td>
<td>.61**</td>
<td>.40**</td>
<td>-.01</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>T: Social</td>
<td>-.13</td>
<td>.46**</td>
<td>.62**</td>
<td>-.02</td>
<td>.52**</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>T: Neurotic</td>
<td>.24**</td>
<td>.10</td>
<td>-.09</td>
<td>.32**</td>
<td>.02</td>
<td>-.28**</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>T: Hist</td>
<td>-.09</td>
<td>.07</td>
<td>.20*</td>
<td>.01</td>
<td>-.17†</td>
<td>-.02</td>
<td>-.04</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>SIT</td>
<td>-.12</td>
<td>.48**</td>
<td>.35**</td>
<td>-.05</td>
<td>.39**</td>
<td>.35**</td>
<td>-.04</td>
<td>.07</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>PID-5-BF</td>
<td>.36**</td>
<td>-.14</td>
<td>-.07</td>
<td>.03</td>
<td>-.19*</td>
<td>-.17†</td>
<td>.14</td>
<td>.14</td>
<td>-.15†</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>IIP-32</td>
<td>.31**</td>
<td>-.11</td>
<td>-.24**</td>
<td>-.02</td>
<td>-.17†</td>
<td>-.28**</td>
<td>.19*</td>
<td>.10</td>
<td>-.09</td>
<td>.57**</td>
<td>—</td>
</tr>
</tbody>
</table>

Table 3. Correlations Among Candidate Predictors and Self-Reported Baseline Depressive Symptoms


**p < .01 *p < .05. †p < .10.
### Table 4. Means on Key Predictors as a Function of Comorbid Anxiety Diagnoses

<table>
<thead>
<tr>
<th></th>
<th>Social Phobia</th>
<th></th>
<th>GAD</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Absent M(SD)</td>
<td>Present M(SD)</td>
<td>Cohen’s d</td>
<td>Absent M(SD)</td>
<td>Present M(SD)</td>
<td>Cohen’s d</td>
<td></td>
<td></td>
</tr>
<tr>
<td>O: Pro-trx</td>
<td>5.00(.88)</td>
<td>4.91(.75)</td>
<td>.11</td>
<td>4.94(.83)</td>
<td>5.04(.84)</td>
<td>-.11</td>
<td></td>
<td></td>
</tr>
<tr>
<td>O: Social</td>
<td>3.33(.90)</td>
<td>3.32(.79)</td>
<td>.00</td>
<td>3.30(.80)</td>
<td>3.40(1.00)</td>
<td>-.11</td>
<td></td>
<td></td>
</tr>
<tr>
<td>O: Neurotic</td>
<td>4.33(.90)</td>
<td>4.47(.87)</td>
<td>-.16</td>
<td>4.40(.88)</td>
<td>4.36(.93)</td>
<td>.04</td>
<td></td>
<td></td>
</tr>
<tr>
<td>T: Pro-trx</td>
<td>4.62(.97)</td>
<td>4.55(1.07)</td>
<td>.07</td>
<td>4.55(.96)</td>
<td>4.71(1.14)</td>
<td>-.14</td>
<td></td>
<td></td>
</tr>
<tr>
<td>T: Social</td>
<td>4.23(1.27)</td>
<td>4.10(1.14)</td>
<td>.10</td>
<td>4.14(1.16)</td>
<td>4.27(1.39)</td>
<td>-.09</td>
<td></td>
<td></td>
</tr>
<tr>
<td>T: Neurotic</td>
<td>4.61(1.12)</td>
<td>4.79(1.10)</td>
<td>-.16</td>
<td>4.71(1.09)</td>
<td>4.62(1.17)</td>
<td>.07</td>
<td></td>
<td></td>
</tr>
<tr>
<td>T: Hist</td>
<td>-.78(.25)</td>
<td>-.74(.30)</td>
<td>-.12</td>
<td>-.77(.27)</td>
<td>-.73(.29)</td>
<td>-.13</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SIT</td>
<td>5.52(.74)</td>
<td>5.02(1.02)</td>
<td>.58**</td>
<td>5.36(.87)</td>
<td>5.11(1.00)</td>
<td>.24</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PID-5-BF</td>
<td>28.04(9.97)</td>
<td>34.22(9.13)</td>
<td>-.64**</td>
<td>29.47(9.95)</td>
<td>34.31(9.65)</td>
<td>-.43*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>IIP-32</td>
<td>48.68(14.50)</td>
<td>58.46(17.15)</td>
<td>-.62**</td>
<td>51.43(15.73)</td>
<td>57.19(17.68)</td>
<td>-.31†</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note. O: Pro-trx = TS-observer: Pro-treatment behaviors; O: Social = TS-observer: Sociability; O: Neurotic = TS-observer: Neuroticism; T: Pro-trx = TS-therapist: Pro-treatment behaviors; T: Social = TS-therapist: Sociability; T: Neurotic = TS-therapist: Neuroticism; T: Hist = TS-therapist: Histrionic; SIT = Standardized Interaction Task; PID-5-BF = Personality Inventory for DSM-5-Brief From; IIP-32 = Inventory of Interpersonal Problems 32-item; GAD = Generalized Anxiety Disorder; M = Mean; SD = Standard Deviation. Significance of mean differences and Cohen’s d Effect Sizes noted from the results of independent sample t-tests. **p < .01 *p < .05. †p < .10.
<table>
<thead>
<tr>
<th></th>
<th>$b^*$</th>
<th>Standard Error</th>
<th>t-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>27.53</td>
<td>0.94</td>
<td>29.21**</td>
</tr>
<tr>
<td>Time</td>
<td>-1.06</td>
<td>0.06</td>
<td>-17.81**</td>
</tr>
<tr>
<td>SIT-observer</td>
<td>0.13</td>
<td>1.10</td>
<td>0.12</td>
</tr>
<tr>
<td>TS-observer: Pro-treatment behaviors</td>
<td>-1.99</td>
<td>1.31</td>
<td>-1.52</td>
</tr>
<tr>
<td>TS-observer: Sociability</td>
<td>0.09</td>
<td>1.27</td>
<td>0.07</td>
</tr>
<tr>
<td>TS-observer: Neuroticism</td>
<td>1.07</td>
<td>0.95</td>
<td>1.12</td>
</tr>
<tr>
<td>IIP-32</td>
<td>1.75</td>
<td>1.20</td>
<td>1.46</td>
</tr>
<tr>
<td>PID-5-BF</td>
<td>1.93</td>
<td>1.20</td>
<td>1.61</td>
</tr>
<tr>
<td>Time*SIT-observer</td>
<td>-0.16</td>
<td>0.07</td>
<td>-2.30*</td>
</tr>
<tr>
<td>Time*TS-observer: Pro-treatment behaviors</td>
<td>0.03</td>
<td>0.08</td>
<td>0.35</td>
</tr>
<tr>
<td>Time*TS-observer: Sociability</td>
<td>0.03</td>
<td>0.08</td>
<td>0.36</td>
</tr>
<tr>
<td>Time*TS-observer: Neuroticism</td>
<td>0.01</td>
<td>0.06</td>
<td>0.14</td>
</tr>
<tr>
<td>Time*IIP-32</td>
<td>0.03</td>
<td>0.08</td>
<td>0.45</td>
</tr>
<tr>
<td>Time*PID-5-BF</td>
<td>-0.20</td>
<td>0.07</td>
<td>-2.68**</td>
</tr>
</tbody>
</table>

Table 5. Parameter Estimates from Mixed Model Examining Effects of Interpersonal Vulnerabilities on Slope of Symptom Change

Note. SIT = Standardized Interaction Task, TS-observer = Thin Slice Observer, IIP-32 = Inventory of Interpersonal Problems 32-item, PID-5-BF = Personality Inventory for DSM-5-Brief From. Time reflects weeks from clients’ intake evaluation. Negative t-values indicate that the given predictor was associated with a more rapid decrease in symptoms from intake to the end of treatment. $b^*$ reflects the beta estimate obtained for a given predictor when all the predictors were first standardized (M = 0, SD =1) prior to being entered in the model.

**$p < .01$ *$p < .05$. † $p < .10$. 
<table>
<thead>
<tr>
<th></th>
<th>Wald χ²</th>
<th>Odds Ratio</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>BDI</td>
<td>0.73</td>
<td>1.29</td>
<td>0.72-2.29</td>
</tr>
<tr>
<td>SIT-observer</td>
<td>0.43</td>
<td>1.22</td>
<td>0.67-2.21</td>
</tr>
<tr>
<td>TS-observer: Pro-treatment behaviors</td>
<td>0.06</td>
<td>0.92</td>
<td>0.46-1.85</td>
</tr>
<tr>
<td>TS-Observer: Sociability</td>
<td>0.55</td>
<td>0.76</td>
<td>0.37-1.57</td>
</tr>
<tr>
<td>TS-Observer: Neuroticism</td>
<td>0.76</td>
<td>0.78</td>
<td>0.45-1.35</td>
</tr>
<tr>
<td>IIP-32</td>
<td>0.12</td>
<td>0.89</td>
<td>0.46-1.72</td>
</tr>
<tr>
<td>PID-5-BF</td>
<td>0.03</td>
<td>1.06</td>
<td>0.56-2.02</td>
</tr>
</tbody>
</table>

Table 6. Observer-rated and Self-reported Interpersonal Vulnerability Predicting Dropout (Controlling for Intake Depressive Symptoms)

Note. SIT = Standardized Interaction Task, TS-observer = Thin Slice Observer, IIP-32 = Inventory of Interpersonal Problems 32-item, PID-5-BF = Personality Inventory for DSM-5-Brief; BDI = Beck Depression Inventory (2nd edition) assessed at clients intake evaluation. All predictors were examined in the same model. For ease of interpretation, predictors were standardized (M = 0, SD = 1).
<table>
<thead>
<tr>
<th>Parameter</th>
<th>( b^* )</th>
<th>Standard Error</th>
<th>t-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>27.77</td>
<td>0.95</td>
<td>29.30**</td>
</tr>
<tr>
<td>Time</td>
<td>-1.07</td>
<td>0.06</td>
<td>-17.20**</td>
</tr>
<tr>
<td>TS-Therapist: Pro-treatment behaviors</td>
<td>-1.08</td>
<td>1.17</td>
<td>-0.93</td>
</tr>
<tr>
<td>TS-Therapist: Sociability</td>
<td>-0.84</td>
<td>1.18</td>
<td>-0.72</td>
</tr>
<tr>
<td>TS-Therapist: Neuroticism</td>
<td>1.89</td>
<td>1.02</td>
<td>1.85†</td>
</tr>
<tr>
<td>TS-Therapist: Histrionic</td>
<td>-1.95</td>
<td>0.96</td>
<td>-2.02*</td>
</tr>
<tr>
<td>Time*TS-Therapist: Pro-treatment behaviors</td>
<td>0.02</td>
<td>0.08</td>
<td>0.23</td>
</tr>
<tr>
<td>Time*TS-Therapist: Sociability</td>
<td>-0.02</td>
<td>0.08</td>
<td>-0.25</td>
</tr>
<tr>
<td>Time*TS-Therapist: Neuroticism</td>
<td>-0.15</td>
<td>0.07</td>
<td>-2.25*</td>
</tr>
<tr>
<td>Time*TS-Therapist: Histrionic</td>
<td>0.04</td>
<td>0.06</td>
<td>0.65</td>
</tr>
</tbody>
</table>

*Table 7. Parameter Estimates from Mixed Model Examining Effects of Therapist Thin Slice Ratings on Slope of Symptom Change*

Note. TS-therapist = Thin Slice Therapist. Time reflects weeks from clients’ intake evaluation. Negative t-values indicate that the given predictor was associated with a more rapid decrease in symptoms from intake to the end of treatment. \( b^* \) reflects the beta estimate obtained for a given predictor when all the predictors were first standardized (M = 0, SD =1) prior to being entered in the model.

**\( p < .01 \) *\( p < .05 \). † \( p < .10 \).
<table>
<thead>
<tr>
<th></th>
<th>Wald χ²</th>
<th>Odds Ratio</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>BDI</td>
<td>0.00</td>
<td>1.00</td>
<td>0.61-1.64</td>
</tr>
<tr>
<td>TS-Therapist: Pro-treatment behaviors</td>
<td>0.77</td>
<td>0.77</td>
<td>0.43-1.38</td>
</tr>
<tr>
<td>TS-Therapist: Sociability</td>
<td>0.24</td>
<td>1.16</td>
<td>0.64-2.10</td>
</tr>
<tr>
<td>TS-Therapist: Neuroticism</td>
<td>0.32</td>
<td>1.17</td>
<td>0.68-1.99</td>
</tr>
<tr>
<td>TS-Therapist: Histrionic</td>
<td>0.09</td>
<td>0.93</td>
<td>0.56-1.53</td>
</tr>
</tbody>
</table>

*Table 8. Therapist Thin Slice Ratings Predicting Dropout (Controlling for Intake Depressive Symptoms)*

Note. TS-therapist = Thin Slice Therapist; BDI = Beck Depression Inventory assessed at clients intake evaluation. All predictors were examined in the same model. For ease of interpretation, predictors were standardized (M = 0, SD = 1).
<table>
<thead>
<tr>
<th></th>
<th>$b^*$</th>
<th>Standard Error</th>
<th>t-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>27.79</td>
<td>0.98</td>
<td>28.38**</td>
</tr>
<tr>
<td>Time</td>
<td>-1.09</td>
<td>0.07</td>
<td>-15.68**</td>
</tr>
<tr>
<td>Pro-treatment behaviors*Rater</td>
<td>-0.05</td>
<td>0.34</td>
<td>-0.16</td>
</tr>
<tr>
<td>Sociability*Rater</td>
<td>-0.05</td>
<td>0.36</td>
<td>-0.15</td>
</tr>
<tr>
<td>Neuroticism*Rater</td>
<td>0.05</td>
<td>0.31</td>
<td>0.17</td>
</tr>
<tr>
<td>Pro-treatment behaviors<em>Rater</em>Time</td>
<td>0.00</td>
<td>0.04</td>
<td>0.11</td>
</tr>
<tr>
<td>Sociability<em>Rater</em>Time</td>
<td>0.01</td>
<td>0.04</td>
<td>0.15</td>
</tr>
<tr>
<td>Neuroticism<em>Rater</em>Time</td>
<td>0.00</td>
<td>0.04</td>
<td>-0.13</td>
</tr>
</tbody>
</table>

*Table 9. Parameter Estimates from Mixed Model Examining Effects of Interactions between Thin Slice Constructs and Rater on Subsequent Symptom Change*

Note. Rater = a binary indicator reflecting whether the value for the specified thin slice construct was obtained from a therapist or observer thin slice rater. Time reflects weeks from clients’ intake evaluation. $b^*$ reflects the beta estimate obtained for a given predictor when all the predictors were first standardized (M = 0, SD = 1) prior to being entered in the model.

**$p < .01 \, \star p < .05. \, \dag p < .10.$**
<table>
<thead>
<tr>
<th></th>
<th>b*</th>
<th>Standard Error</th>
<th>t-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>58.78</td>
<td>0.94</td>
<td>62.85**</td>
</tr>
<tr>
<td>Time</td>
<td>0.61</td>
<td>0.23</td>
<td>2.68**</td>
</tr>
<tr>
<td>BDI-Intake</td>
<td>0.21</td>
<td>0.90</td>
<td>0.23</td>
</tr>
<tr>
<td>SIT-observer</td>
<td>0.23</td>
<td>1.08</td>
<td>0.22</td>
</tr>
<tr>
<td>TS-observer: Pro-treatment behaviors</td>
<td>2.75</td>
<td>1.29</td>
<td>2.13*</td>
</tr>
<tr>
<td>TS-observer: Sociability</td>
<td>0.11</td>
<td>1.22</td>
<td>0.09</td>
</tr>
<tr>
<td>TS-observer: Neuroticism</td>
<td>1.06</td>
<td>0.98</td>
<td>1.08</td>
</tr>
<tr>
<td>SIT-observer*Time</td>
<td>0.22</td>
<td>0.26</td>
<td>0.83</td>
</tr>
<tr>
<td>TS-observer: Pro-treatment behaviors*Time</td>
<td>-0.27</td>
<td>0.31</td>
<td>-0.87</td>
</tr>
<tr>
<td>TS-observer: Sociability*Time</td>
<td>0.14</td>
<td>0.30</td>
<td>0.48</td>
</tr>
<tr>
<td>TS-observer: Neuroticism*Time</td>
<td>0.00</td>
<td>0.24</td>
<td>0.00</td>
</tr>
</tbody>
</table>

Table 10. Parameter Estimates from Mixed Model Examining Effects of Observer-Rated Interpersonal Vulnerability on Average Alliance Scores and Slope of Alliance

Note. BDI = Beck Depression Inventory (2nd edition), assessed at clients’ intake evaluation; SIT-observer = Standardized Interaction Task; TS-observer = Thin Slice Observer Measure. All predictors were examined in the same model. The time variable was person-mean centered so that the Alliance-intercept predicted for each client would reflect clients’ average Alliance score. For ease of interpretation, predictors were standardized (M = 0, SD = 1).

**p < .01 *p < .05. † p < .10.
<table>
<thead>
<tr>
<th></th>
<th>b*</th>
<th>Standard Error</th>
<th>t-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>59.12</td>
<td>0.87</td>
<td>67.60**</td>
</tr>
<tr>
<td>Time</td>
<td>0.52</td>
<td>0.21</td>
<td>2.49*</td>
</tr>
<tr>
<td>BDI</td>
<td>-1.54</td>
<td>0.48</td>
<td>-3.19**</td>
</tr>
<tr>
<td>BDI*Time</td>
<td>-0.26</td>
<td>0.20</td>
<td>-1.33</td>
</tr>
<tr>
<td>TS-observer: Pro-treatment behaviors</td>
<td>2.40</td>
<td>0.89</td>
<td>2.69**</td>
</tr>
<tr>
<td>PID-5-BF</td>
<td>1.42</td>
<td>1.09</td>
<td>1.30</td>
</tr>
<tr>
<td>IIP-32</td>
<td>-1.80</td>
<td>1.08</td>
<td>-1.67</td>
</tr>
<tr>
<td>TS-observer: Pro-treatment behaviors*Time</td>
<td>0.09</td>
<td>0.21</td>
<td>-0.43</td>
</tr>
<tr>
<td>PID-5-BF*Time</td>
<td>0.46</td>
<td>0.25</td>
<td>1.84†</td>
</tr>
<tr>
<td>IIP-32*Time</td>
<td>-0.30</td>
<td>0.25</td>
<td>-1.21</td>
</tr>
</tbody>
</table>

*Table 11. Parameter Estimates from Mixed Model Examining Effects of Thin Slice Observer: Pro-treatment Behaviors on Alliance while Controlling for Potential Confounds*

Note. BDI = Beck Depression Inventory (2nd edition), assessed at each of the first five therapy sessions; TS-observer = Thin Slice Observer Measure; IIP-32 = Inventory of Interpersonal Problems 32-item, PID-5-BF = Personality Inventory for DSM-5-Brief From. All predictors were examined in the same model. The time variable was person-mean centered so that the WAI-intercept predicted for each client would reflect clients’ average Alliance score. For ease of interpretation, predictors were standardized (M = 0, SD = 1).

**p < .01 *p < .05. † p < .10.
<table>
<thead>
<tr>
<th></th>
<th>( b^* )</th>
<th>Standard Error</th>
<th>t-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>15.45</td>
<td>0.45</td>
<td>34.22**</td>
</tr>
<tr>
<td>Time</td>
<td>0.51</td>
<td>0.13</td>
<td>3.79**</td>
</tr>
<tr>
<td>BDI-Intake</td>
<td>0.63</td>
<td>0.46</td>
<td>1.38</td>
</tr>
<tr>
<td>SIT-observer</td>
<td>-0.49</td>
<td>0.52</td>
<td>-0.94</td>
</tr>
<tr>
<td>TS-observer: Pro-treatment behaviors</td>
<td>1.69</td>
<td>0.62</td>
<td>2.72**</td>
</tr>
<tr>
<td>TS-observer: Sociability</td>
<td>0.40</td>
<td>0.59</td>
<td>0.68</td>
</tr>
<tr>
<td>TS-observer: Neuroticism</td>
<td>0.56</td>
<td>0.47</td>
<td>1.19</td>
</tr>
<tr>
<td>SIT-observer*Time</td>
<td>-0.02</td>
<td>0.15</td>
<td>-0.15</td>
</tr>
<tr>
<td>TS-observer: Pro-treatment behaviors*Time</td>
<td>-0.16</td>
<td>0.18</td>
<td>-0.85</td>
</tr>
<tr>
<td>TS-observer: Sociability*Time</td>
<td>0.09</td>
<td>0.17</td>
<td>0.49</td>
</tr>
<tr>
<td>TS-observer: Neuroticism*Time</td>
<td>-0.04</td>
<td>0.14</td>
<td>-0.29</td>
</tr>
</tbody>
</table>

Table 12. Parameter Estimates from Mixed Model Examining Effects of Observer-Rated Interpersonal Vulnerability on Average Cognitive Change and Slope of Cognitive Change

Note. BDI = Beck Depression Inventory, assessed at intake evaluation (2nd edition); SIT-observer = Standardized Interaction Task; TS-observer = Thin Slice Observer Measure. All predictors were examined in the same model. The time variable was person-mean centered so that the Cognitive Change Score-intercept predicted for each client would reflect clients’ average in-session cognitive change score. For ease of interpretation, predictors were standardized (\( M = 0, SD = 1 \)).

**\( p < .01 \) *\( p < .05 \) †\( p < .10 \).
Table 13. Parameter Estimates from Mixed Model Examining Effects of Thin Slice Observer: Pro-treatment Behaviors on Cognitive Change while Controlling for Potential Confounds

Note. BDI = Beck Depression Inventory (2nd edition), assessed at each of the first five therapy sessions; TS-observer = Thin Slice Observer; IIP-32 = Inventory of Interpersonal Problems 32-item; PID-5-BF = Personality Inventory for DSM-5-Brief Form. All predictors were examined in the same model. The time variable was person-mean centered so that the intercept of Cognitive Change (CC) scores predicted for each client would reflect clients’ average CC score. For ease of interpretation, predictors were standardized (M = 0, SD = 1).

**p < .01 *p < .05 †p < .10.

<table>
<thead>
<tr>
<th></th>
<th>b*</th>
<th>Standard Error</th>
<th>t-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>15.43</td>
<td>0.44</td>
<td>35.13**</td>
</tr>
<tr>
<td>Time</td>
<td>0.44</td>
<td>0.13</td>
<td>3.30**</td>
</tr>
<tr>
<td>BDI</td>
<td>-0.51</td>
<td>0.33</td>
<td>-1.52</td>
</tr>
<tr>
<td>BDI*Time</td>
<td>-0.19</td>
<td>0.14</td>
<td>-1.39</td>
</tr>
<tr>
<td>TS-observer: Pro-treatment behaviors</td>
<td>1.46</td>
<td>0.45</td>
<td>3.26**</td>
</tr>
<tr>
<td>PID-5-BF</td>
<td>1.15</td>
<td>0.55</td>
<td>2.11*</td>
</tr>
<tr>
<td>IIP-32</td>
<td>-0.84</td>
<td>0.54</td>
<td>-1.54</td>
</tr>
<tr>
<td>TS-observer: Pro-treatment behaviors*Time</td>
<td>-0.15</td>
<td>0.13</td>
<td>-1.16</td>
</tr>
<tr>
<td>PID-5-BF*Time</td>
<td>-0.07</td>
<td>0.16</td>
<td>-0.46</td>
</tr>
<tr>
<td>IIP-32*Time</td>
<td>0.04</td>
<td>0.16</td>
<td>0.22</td>
</tr>
<tr>
<td></td>
<td>b*</td>
<td>Standard Error</td>
<td>t-value</td>
</tr>
<tr>
<td>--------------------------</td>
<td>---------</td>
<td>----------------</td>
<td>----------</td>
</tr>
<tr>
<td>Intercept</td>
<td>58.86</td>
<td>0.90</td>
<td>65.48**</td>
</tr>
<tr>
<td>Time</td>
<td>0.58</td>
<td>0.22</td>
<td>2.66**</td>
</tr>
<tr>
<td>BDI-Intake</td>
<td>0.41</td>
<td>0.87</td>
<td>0.47</td>
</tr>
<tr>
<td>TS-therapist: Histrionic</td>
<td>0.21</td>
<td>0.92</td>
<td>0.23</td>
</tr>
<tr>
<td>TS-therapist: Pro-treatment behaviors</td>
<td>1.75</td>
<td>1.10</td>
<td>1.60</td>
</tr>
<tr>
<td>TS-therapist: Sociability</td>
<td>2.28</td>
<td>1.13</td>
<td>2.02*</td>
</tr>
<tr>
<td>TS-therapist: Neuroticism</td>
<td>0.14</td>
<td>0.98</td>
<td>0.14</td>
</tr>
<tr>
<td>TS-therapist: Histrionic*Time</td>
<td>-0.13</td>
<td>0.22</td>
<td>-0.58</td>
</tr>
<tr>
<td>TS-therapist: Pro-treatment behaviors*Time</td>
<td>-0.27</td>
<td>0.27</td>
<td>-0.99</td>
</tr>
<tr>
<td>TS-therapist: Sociability*Time</td>
<td>0.26</td>
<td>0.28</td>
<td>0.92</td>
</tr>
<tr>
<td>TS-therapist: Neuroticism*Time</td>
<td>-0.26</td>
<td>0.23</td>
<td>-1.13</td>
</tr>
</tbody>
</table>

*Table 14. Parameter Estimates from Mixed Model Examining Effects of Therapist-Rated Interpersonal Vulnerability on Average Alliance Scores and Slope of Alliance*

Note. BDI = Beck Depression Inventory, assessed at intake evaluation (2nd edition); TS-therapist = Thin Slice Therapist. All predictors were examined in the same model. The time variable was person-mean centered so that the Alliance-intercept predicted for each client would reflect clients’ average Alliance score. For ease of interpretation, predictors were standardized (M = 0, SD = 1).

**p < .01  *p < .05  † p < .10.
<table>
<thead>
<tr>
<th></th>
<th>$b^*$</th>
<th>Standard Error</th>
<th>t-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>59.08</td>
<td>0.87</td>
<td>67.76**</td>
</tr>
<tr>
<td>Time</td>
<td>0.45</td>
<td>0.21</td>
<td>2.16*</td>
</tr>
<tr>
<td>BDI</td>
<td>-1.77</td>
<td>0.48</td>
<td>-3.73**</td>
</tr>
<tr>
<td>BDI*Time</td>
<td>-0.26</td>
<td>0.20</td>
<td>-1.33</td>
</tr>
<tr>
<td>TS-therapist: Sociability</td>
<td>2.80</td>
<td>0.91</td>
<td>3.06**</td>
</tr>
<tr>
<td>PID-5-BF</td>
<td>0.85</td>
<td>1.07</td>
<td>0.80</td>
</tr>
<tr>
<td>IIP-32</td>
<td>-0.81</td>
<td>1.09</td>
<td>-0.74</td>
</tr>
<tr>
<td>TS-therapist: Sociability*Time</td>
<td>0.14</td>
<td>0.21</td>
<td>0.65</td>
</tr>
<tr>
<td>PID-5-BF*Time</td>
<td>0.37</td>
<td>0.25</td>
<td>1.49</td>
</tr>
<tr>
<td>IIP-32*Time</td>
<td>-0.16</td>
<td>0.25</td>
<td>-0.64</td>
</tr>
</tbody>
</table>

Table 15. Parameter Estimates from Mixed Model Examining Effects of Thin Slice Therapist: Sociability on Alliance while Controlling for Potential Confounds

Note. BDI = Beck Depression Inventory (2nd edition), assessed at each of the first five therapy sessions; TS-therapist = Thin Slice Therapist; IIP-32 = Inventory of Interpersonal Problems 32-item, PID-5-BF = Personality Inventory for DSM-5-Brief From. All predictors were examined in the same model. The time variable was person-mean centered so that the Alliance-intercept predicted for each client would reflect clients’ average Alliance score. For ease of interpretation, predictors were standardized (M = 0, SD = 1).

**$p < .01$ *$p < .05$. †$p < .10$. 

120
<table>
<thead>
<tr>
<th></th>
<th>$b^*$</th>
<th>Standard Error</th>
<th>t-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>15.54</td>
<td>0.45</td>
<td>34.33**</td>
</tr>
<tr>
<td>Time</td>
<td>0.48</td>
<td>0.13</td>
<td>3.66**</td>
</tr>
<tr>
<td>BDI-Intake</td>
<td>0.62</td>
<td>0.46</td>
<td>1.33</td>
</tr>
<tr>
<td>TS-therapist: Histrionic</td>
<td>0.60</td>
<td>0.47</td>
<td>1.28</td>
</tr>
<tr>
<td>TS-therapist: Pro-treatment</td>
<td>0.70</td>
<td>0.55</td>
<td>1.26</td>
</tr>
<tr>
<td>TS-therapist: Sociability</td>
<td>0.83</td>
<td>0.57</td>
<td>1.45</td>
</tr>
<tr>
<td>TS-therapist: Neurotic</td>
<td>0.25</td>
<td>0.50</td>
<td>0.50</td>
</tr>
<tr>
<td>TS-therapist: Histrionic*Time</td>
<td>-0.04</td>
<td>0.14</td>
<td>-0.29</td>
</tr>
<tr>
<td>TS-therapist: Pro-treatment*Time</td>
<td>-0.14</td>
<td>0.16</td>
<td>-0.87</td>
</tr>
<tr>
<td>TS-therapist: Sociability*Time</td>
<td>0.00</td>
<td>0.16</td>
<td>0.00</td>
</tr>
<tr>
<td>TS-therapist: Neurotic*Time</td>
<td>-0.01</td>
<td>0.14</td>
<td>-0.08</td>
</tr>
</tbody>
</table>

Table 16. Parameter Estimates from Mixed Model Examining Effects of Therapist-Rated Interpersonal Vulnerability on Average In-session Cognitive Change Scores and Slope of In-session Cognitive Change

Note. BDI = Beck Depression Inventory, assessed at intake evaluation (2nd edition); TS-therapist = Thin Slice Therapist. All predictors were examined in the same model. The time variable was person-mean centered so that the in-session cognitive change (CC) scores predicted for each client would reflect the clients’ average CC score. For ease of interpretation, predictors were standardized ($M = 0, SD = 1$).

** $p < .01$  * $p < .05$. † $p < .10$. 
<table>
<thead>
<tr>
<th></th>
<th>$b^*$</th>
<th>Standard Error</th>
<th>t-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>62.16</td>
<td>1.18</td>
<td>52.74**</td>
</tr>
<tr>
<td>Time</td>
<td>1.07</td>
<td>0.38</td>
<td>2.82**</td>
</tr>
<tr>
<td>BDI</td>
<td>-0.13</td>
<td>0.03</td>
<td>-4.57**</td>
</tr>
<tr>
<td>BDI*Time</td>
<td>-0.02</td>
<td>0.01</td>
<td>-1.64</td>
</tr>
<tr>
<td>Pro-treatment behaviors*Rater</td>
<td>0.10</td>
<td>0.29</td>
<td>0.34</td>
</tr>
<tr>
<td>Sociability*Rater</td>
<td>0.00</td>
<td>0.30</td>
<td>0.01</td>
</tr>
<tr>
<td>Neuroticism*Rater</td>
<td>0.06</td>
<td>0.26</td>
<td>0.22</td>
</tr>
<tr>
<td>Pro-treatment behaviors<em>Rater</em>Time</td>
<td>-0.10</td>
<td>0.19</td>
<td>-0.55</td>
</tr>
<tr>
<td>Sociability<em>Rater</em>Time</td>
<td>0.05</td>
<td>0.20</td>
<td>0.24</td>
</tr>
<tr>
<td>Neuroticism<em>Rater</em>Time</td>
<td>-0.02</td>
<td>0.17</td>
<td>-0.15</td>
</tr>
</tbody>
</table>

Table 17. Parameter Estimates from Mixed Model Examining Effects of Interactions between Thin Slice Constructs and Rater on Alliance

Note. BDI = Beck Depression Inventory (2nd edition), assessed at each of the first five therapy sessions; Rater = a binary indicator reflecting whether the values for the specified thin slice construct were obtained from a therapist or observer thin slice rater. The time variable was person-mean centered so that the alliance scores predicted for each client would reflect clients’ average Alliance score. $b^*$ reflects the beta estimate obtained for a given predictor when all the predictors were first standardized (M = 0, SD = 1) prior to being entered in the model.

**$p < .01$ *$p < .05$. †$p < .10$. **
<table>
<thead>
<tr>
<th></th>
<th>( b^* )</th>
<th>Standard Error</th>
<th>t-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>16.63</td>
<td>0.72</td>
<td>23.05**</td>
</tr>
<tr>
<td>Time</td>
<td>0.46</td>
<td>0.30</td>
<td>1.55</td>
</tr>
<tr>
<td>BDI</td>
<td>-0.04</td>
<td>0.02</td>
<td>-2.02*</td>
</tr>
<tr>
<td>BDI*Time</td>
<td>0.00</td>
<td>0.01</td>
<td>-0.09</td>
</tr>
<tr>
<td>Pro-treatment behaviors*Rater</td>
<td>0.12</td>
<td>0.24</td>
<td>0.51</td>
</tr>
<tr>
<td>Sociability*Rater</td>
<td>0.04</td>
<td>0.25</td>
<td>0.17</td>
</tr>
<tr>
<td>Neuroticism*Rater</td>
<td>0.07</td>
<td>0.22</td>
<td>0.33</td>
</tr>
<tr>
<td>Pro-treatment behaviors<em>Rater</em>Time</td>
<td>-0.07</td>
<td>0.16</td>
<td>-0.48</td>
</tr>
<tr>
<td>Sociability<em>Rater</em>Time</td>
<td>0.00</td>
<td>0.17</td>
<td>0.01</td>
</tr>
<tr>
<td>Neuroticism<em>Rater</em>Time</td>
<td>-0.05</td>
<td>0.14</td>
<td>-0.37</td>
</tr>
</tbody>
</table>

Table 18. Parameter Estimates from Mixed Model Examining Effects of Interactions between Thin Slice Constructs and Rater on In-session Cognitive Change

Note. BDI = Beck Depression Inventory (2nd edition), assessed at each of the first five therapy sessions; Rater = a binary indicator reflecting whether the values for the specified thin slice construct was obtained from a therapist or observer thin slice rater. The time variable was person-mean centered so that the in-session cognitive change (CC) scores predicted for each client would reflect clients’ average CC score. \( b^* \) reflects the beta estimate obtained for a given predictor when all the predictors were first standardized (M = 0, SD = 1) prior to being entered in the model. **\( p < .01 \) *\( p < .05 \). † \( p < .10 \).
Appendix B: Figures
Note. Parallel analysis suggests three dimensions comprise the Thin Slice (TS) Observer measure, as the dashed red line (plotting the median eigenvalues observed across $k = 1,000$ randomly generated correlation matrices of the same dimension as the TS-observer data) crosses the solid factor analyses line (plotting observed eigenvalues) before reaching the fourth factor.

*Figure 1. Scree Plot of Eigenvalues Derived from Parallel Analysis of Thin Slice Observer Measure*
Figure 2. Scree Plot of Eigenvalues Derived from Parallel Analysis of Thin Slice Therapist Measure

Note. Parallel analysis suggests four dimensions comprise the Thin Slice (TS) Therapist measure, as the dashed line (plotting the median eigenvalues observed across $k = 1,000$ randomly generated correlation matrices of the same dimension as TS-therapist data) crosses the solid factor analyses line (plotting observed eigenvalues) before reaching the fifth factor.
Figure 3. Scree Plot of Eigenvalues Derived from Parallel Analysis of In-Session Cognitive Change Measure

Note. Parallel analysis suggests one dimension comprises the In-Session Cognitive Change (CC) Measure, as the dashed line (plotting the median eigenvalues observed across $k = 1,000$ randomly generated correlation matrices of the same dimension as CC data) crosses the solid factor analyses line (plotting observed eigenvalues) before reaching the second factor.