Assessing Patients' Mastery of the Skills of Cognitive Therapy: Initial Evaluation of the Patient Competencies Scale

THESIS

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

By

Shannon Nicole Hollars

Graduate Program in Psychology

The Ohio State University

2012

Master's Examination Committee:

Dr. Daniel R. Strunk, Advisor

Dr. Jennifer S. Cheavens

Dr. Jennifer K. Crocker

Copyrighted by

Shannon Nicole Hollars

2012

Abstract

In Cognitive Therapy (CT), therapists work to help patients develop a specific set of skills to cope with negative emotions. While mastery and use of these skills has been related to reduced risk for relapse, current methods of assessing patients' mastery of these skills are cumbersome and would present substantial difficulties for clinical use. In this study, we sought to examine the validity of a brief and easily scored measure of CT skills and investigated if therapist techniques might be a more robust predictor of symptom reduction in early sessions for patients who possess greater pre-treatment skills. In a sample of 65 depressed patients participating in CT, we pilot test the new Patient Competencies Scale (PCS)—with one version to be completed by patients and a second version by therapists. The PCS correlated with intake and post treatment ratings on related measures of skills and depression severity as well as change in these assessments over the course of treatment. Client rated skill acquisition predicted reductions in depressive symptoms over the course of therapy even when including scores from the Ways of Responding questionnaire as a covariate. Neither patient nor therapist versions of the PCS was found to be predictive of risk of relapse in the year following acute treatment. We compare the PCS to previous self-report measures and recommend the development of a clinical interview to assess patient competencies. Furthermore, because therapists' use of Socratic questioning has been posited to foster patients' independent usage of these skills, we examine whether therapists' use of these techniques interacted with patients' baseline skills in predicting early symptom change. We report on an interaction between baseline skills and Socratic questioning in predicting early symptom

ii

change. For patients with higher baseline skills, Socratic questioning was particularly strongly related to early session-to-session symptom improvements.

Acknowledgments

First, I would like to extend my gratitude to Dr. Daniel Strunk for serving as a patient and encouraging mentor throughout this process. His guidance has been invaluable in helping me determine which questions are worth pursuing and how to answer them using appropriate methods and analyses. I am thankful for his expertise and commitment both to this project and my own development as a researcher. I also want to express my gratitude to the members of my master's thesis committee. Their feedback has informed my conceptualization of my study and deepened my understanding of this field of research. I am also immensely grateful to my parents who have always encouraged and supported me as I pursue my goals. Their influence has been nothing short of instrumental to my personal, academic, and professional development. Finally, I would like to express my appreciation for the support and advice of my fellow graduate students in the Depression Research Lab. These analyses would not have been possible without their care and dedication in collecting these data.

Vita

June 1987	Born, Charlotte, North Carolina
2009	B.A. in Psychology with highest honors,
	Summa Cum Laude, Emory University
2010 to present	Graduate Student and Graduate Teaching
	Assistant, Department of Psychology, The
	Ohio State University

Fields of Study

Major Field: Psychology

Minor Field: Quantitative Psychology

Table of Contents

Abstractii
Acknowledgmentsiv
Vitav
List of Tables
Chapter 1: Introduction
Variability in Outcome Achieved with CT
Socratic Questioning
Patient Skill Development in CT9
Assessment of Patient Skills 10
Skills of Cognitive Therapy Scale
Purpose
Chapter 2: Method
Participants 16
Measures17
Axis I Diagnosis
Depressive Symptom Severity17
Patient Competencies in CT Skills

Depression Relapse	20
Socratic Questioning	20
Procedures	21
Chapter 3: Results	22
Validation analyses for the PCS-C and PCS-T	22
The Factor Structure of the PCS	22
Detecting Change from Pre to Post Treatment	23
Differentiating between Patients with MDD and Controls	24
Correlations with Depression Severity and Related Skills Measures	24
Predicting Change in Depression Severity	27
Predicting Risk of Relapse	31
Socratic Questioning Analyses	32
Chapter 4: Discussion	35
Limitations	42
References	45
Appendix A: Patient Competencies Scale- Client Version	51
Appendix B: Patient Competencies Scale- Therapist Version	55
Appendix C: Tables	57

List of Tables

Table 1: Correlations of pre and post treatment patient	
competencies with measures of CT skills and depression	
symptom severity	

Table 2: Correlations of PCS residualized change with pre	
to post treatment change in alternative compensatory skills	
measures and slopes of change of depressive severity	58

Table 3: Parameter estimates (and Standard Errors) for Mixed	
Models examining residualized change in WOR Quality and	
PCS as predictors of HRSD and BDI-II change	59

Table 4: Parameter estimates (and Standard Errors) for Mixed	
Models examining residualized change in WOR Total and PCS	
as predictors of HRSD and BDI-II change	60

Chapter 1: Introduction

Cognitive Therapy (CT) is among the most studied psychotherapies for depression. As an acute treatment, previous trials have shown CT to be as efficacious as antidepressant medication (ADM; Strunk & DeRubeis, 2001) and superior to pill-placebo (Jarrett et al., 1999; DeRubeis et al., 2005). CT also appears to have enduring effects. In a large study of moderate to severe depression, patients who had participated in a course of CT exhibited a comparable risk of relapse to those continued on ADM and a significantly lower risk of relapse than those who discontinued ADM (Hollon et al., 2005). In fact, there is evidence of such enduring effects for CT across depression and other emotional disorders (Hollon, Stewart, & Strunk, 2006). What accounts for the enduring effects of CT is not well understood. Cognitive and behavioral skills acquired in the course of CT have been proposed as an important mechanism through which CT may achieve such enduring results (Beck, 1967, 1976). Accordingly, helping patients develop these CT skills is a primary focus of the therapeutic process in CT. Therapists are thought to foster patients' independent application of these skills outside of therapy through the use of Socratic questioning, a collaborative process of guided self-discovery. While Socratic questioning has been highlighted as a requirement for the competent implementation of CT, it is also understood to be a particularly challenging skill for therapists to master (Beck, Rush, Shaw, & Emery, 1979; IAPT Programme, 2007; Overholser, 2011). Differences in therapists' competent implementation of Socratic techniques make it a logical source of variability observed in patient outcomes in CT.

Over the course of CT treatment, therapists seek to teach patients a set of CTspecific cognitive and behavioral skills that they can continue to utilize after ending treatment (Beck et al., 1979). Previous research using a sample of moderate to severely depressed patients who responded to 16 weeks of acute CT treatment found that greater post treatment CT coping skills predicted reduced risk of relapse in the one year following treatment, even after controlling for the influence of pretreatment skill level, post treatment symptom severity, and change in symptoms over the course of treatment (Strunk, DeRubeis, Chiu, & Alvarez, 2007). In this study by Strunk and colleagues (2007) as well as in multiple related studies of CT coping skills, the Ways of Responding (WOR) questionnaire (Barber & DeRubeis, 1992) has been used as the primary measure of patients' acquisition of CT. This measure requires trained raters to score patient responses to events and accompanying cognitions in terms of the level of CT coping skills. Consequently, this and other currently available measures to assess these skills involve a lengthy rating process that makes them impractical for clinical practice. A brief and readily scored measure of these skills would have obvious benefit for clinicians.

Given the potential role of coping skills in the efficacy of CT for Major Depressive Disorder (MDD) and the challenge of getting patients to develop these skills in CT, we sought to investigate how therapist techniques might relate to patients' competencies in coping skills. Specifically, since Socratic questioning has been posited to be a requirement for the competent implementation of CT, we examined whether therapists' use of this strategy interacted with patients' pre-treatment competencies in these skills to predict early treatment outcomes. Moreover, because the labor-intensive rating process to assess these skills makes them impractical to measure systematically in clinical practice, developing and validating a measure to assess these skills efficiently was the second aim of this research. The following sections will review the relevant literature and research on Socratic questioning and coping skills, highlighting methods whereby this technique has been proposed to impact the efficacy of CT, and examining the role of patient competencies in fostering and maintaining depressive symptom change. Finally, the rationale for our development and validation of a brief self-report and therapist rated measure of skills will be introduced.

Variability in Outcome Achieved with CT

While overall results for CT have established its efficacy, the performance of CT has been variable, with outcomes in some studies (Dimidjian et al., 2006) or at sites within multi-site clinical trials (Elkin et al., 1995; DeRubeis et al., 2005) being lower than expected based on outcomes from similar studies. Variability across studies in the performance of CT compared to other treatments, namely ADM, has produced conflicting findings concerning the comparative efficacy of these treatments and led to questions about potential sources of this variability. In Dimidijian and colleagues' (2006) randomized, placebo controlled clinical trial of adults with MDD, severely depressed patients showed improvement with ADM, and ADM performed significantly better than CT. Moreover, early doubt about the relative efficacy of ADM and CT was raised after the analyses of the Treatment of Depression Collaborative Research Program Study (TDCRP; Elkin et al., 1989), a multi-site, randomized clinical trial among adult outpatients with MDD. Follow-up analyses of these data using severity as a moderator

found no significant differences between ADM and CT, but only ADM (Imipramine plus clinical management) and not CT outperformed placebo with clinical management among severely depressed patients (Elkin et al., 1995).

While follow-up analyses of the TDCRP findings using severity as a moderator failed to establish CT as comparably efficacious to ADM in the treatment of moderate to severe MDD (Elkin et al., 1995), multiple studies have produced contrasting results showing CT to produce equivalent results to medication. A mega-analysis pooling data from four randomized clinical trials (including TDCRP) that compared ADM and CT for MDD failed to find support for the conclusion that ADM is superior to CT in the treatment of severe depression (DeRubeis, Gelfand, Tang, & Simons, 1999). In addition, DeRubeis and colleagues (2005) conducted a subsequent two-site randomized clinical trial comparing CT and ADM in the treatment of moderate to severe depression and failed to find a significant difference between CT and ADM. Moreover, both active treatments showed significant differences in symptom reduction when compared to placebo.

Because treatment procedures are likely to be more similar within a single study than between separate studies, multi-site trials that report variability in the performance of CT across sites may present a greater complication for understanding the comparative efficacy of CT than finding variable performance between separate trials with unique protocols and patient samples. In an effort to account for such site by treatment interactions, discrepancies in treatment implementation and differences in patient characteristics have been proposed to be potential sources of this variability. In the TDCRP findings, for example, CT performed significantly better than placebo and similarly to ADM at one of the three sites. While the primary investigators have never released which results originated from which site, it is notable that one of the sites is known to have a strong CT focus while another is the location of the first research conducted on IPT (Elkin et al., 1995). Therefore, the variable results in treatment efficacy across these sites have been posited to have resulted from differences in therapist expertise and the implementation of CT. In support of this supposition, DeRubeis and colleagues' (2005) found a significant site by treatment interaction such that ADM was found to outperform CT at Vanderbilt University where CT therapists had comparably less experience with CT, and CT did not perform significantly better than ADM at the University of Pennsylvania site. The authors argued that the existence of site differences in the relative experience of CT therapists may have led to differences in the implementation of CT and influenced the efficacy of the treatment. However, the authors also reported significant differences in patient characteristics between sites that could also explain some of the variability in the performance of CT in their two-site clinical trial. Notably, at Vanderbilt University where CT was comparatively less efficacious, patients were more likely to be female, have a diagnosis of Post-traumatic stress disorder, and have a comorbid Axis I disorder (DeRubeis et al., 2005). While it has been difficult to determine the relative influence of patient characteristics and therapist experience on the variability in CT performance in this trial, such evidence implicates these two factors as plausible sources of the variability in the performance of CT both across studies and within multi-site trials.

5

Socratic Questioning

Because differences in the way CT has been implemented are often discussed as potential sources of variability in outcome, therapist use of Socratic techniques has been targeted as a possible domain whereby therapists may differ in implementing CT. Socratic questioning is a central aspect of conducting CT which experts have suggested may be difficult for therapists to master (Beck, 1979; Overholser, 2011). This technique refers to a therapist's use of a series of targeted questions aimed at teaching patients to identify and evaluate their negative automatic thoughts independently and, ultimately, without requiring therapist aid. As treatment progresses, Socratic questioning might be used as a key part of a therapist's effort to help patients autonomously use the skills of CT. Because this process is regarded as requiring greater therapist expertise, it is a logical candidate for explaining the variability in the efficacy of CT observed both within and between studies (Beck, 1979; Overholser, 2011).

The model underlying CT posits that depressive symptoms are maintained by patient's distorted negative beliefs about themselves, the world, and the future (Beck 1967, 1976). Such distorted beliefs, referred to as dysfunctional attitudes, are thought to result from early life experiences that result in underlying schemas usually involving personal failure, loss, worthlessness, and inadequacy. These dysfunctional attitudes are thought to lead to negative automatic thoughts, or spontaneous cognitions in response to a situation. CT targets these dysfunctional attitudes and corresponding automatic negative thoughts as a primary method for enacting symptom change.

In CT, therapists aim to impart in patients a set of cognitive and behavioral skills through a collaborative process of guided self-discovery. This Socratic process relies on a therapist's use of a series of questions designed to help patients identify their own automatic negative thoughts and question their validity. This systematic questioning is intended to be a collaborative effort between therapist and patient whereby the therapist tactfully helps the patient to identify errors in his or her thinking. The collaborative nature of Socratic questioning provides patients with practice examining evidence for their automatic thoughts, generating alternative explanations, and reevaluating their beliefs. This process of guided self-discovery is thought to increase client motivation to reevaluate their thought processes and promote their independent, competent use of these skills outside of therapy. The Socratic method has been posited to facilitate change by promoting patient autonomy and fostering an internal locus of change even early in the course of treatment (Overholser, 2011). Early in treatment, patients who have greater initial mastery of CT skills might be expected to show greater symptom improvements when therapists utilize Socratic techniques. This expectation is based on the assumption that Socratic techniques require patient contributions, and those patients with greater skills are more likely to identify useful cognitive strategies which could facilitate symptom improvement (Overholser, 2011).

Despite the theorized and frequently espoused importance of these techniques to CT, scant research has been conducted to examine if Socratic questioning is indeed related to skill acquisition or patient outcomes. Nonetheless, a task force initiative assembled by the British Psychological Society's Centre for Outcomes Research and Effectiveness that was charged with developing a list of therapist skills required to conduct CT properly has pronounced the Socratic method as necessary for the competent administration of CT (IAPT Programme, 2007). To our knowledge, only a single study has examined the association between Socratic questioning and symptom change in CT for MDD (Braun & Strunk, personal communication). In a study of CT for adult outpatients with MDD, observer rated questions assessing novice therapist use of the Socratic method were found to predict session-to-session symptom change in the first three sessions of CT and to differentiate novice and expert therapists. These preliminary findings provide initial support for the relationship between the Socratic method and early treatment outcomes as well as the posited expertise required to implement this technique.

Though we are not aware of an empirical test of this issue, we would expect Socratic questioning to be more useful for patients who already demonstrate competence in the use of CT skills. These patients' greater facility with these skills would likely enable them to respond more adaptively to early efforts at Socratic questioning. Patients with higher pre-treatment skill levels may be initially more familiar with and open to the central tenets of CT, and therefore may be more responsive to a therapist's collaborative, Socratic approach. Rather than having to spend considerable time covering the rationale for treatment and discussing some of the more basic aspects of CT, therapists could improve patients' already existing skills and elicit skill use likely to impact moods more quickly. Should such an interaction between pre-treatment skills and Socratic techniques be present, therapists who could identify patients with greater coping skills at the onset of treatment could have the potential to maximize patient outcomes. However, currently available measures of these skills are too labor intensive and time consuming to be practical for clinical use, so a valid, short skills assessment would be necessary for therapists to exploit any relationship between skills and Socratic questioning.

Patient Skill Development in CT

CT skills are thought to allow patients to address stressors in more adaptive ways by helping them cope with their negative thoughts and moods. The examination of evidence for automatic thoughts is a central focus of the cognitive skills that patients build in CT. Patients are taught first to recognize the connection between their thoughts and mood, and to identify and then assess critically the accuracy of their automatic thoughts. As part of this process, patients learn to examine the evidence for and against their thoughts, to conduct behavioral experiments to test their negative beliefs, to generate potential alternative explanations, and to reevaluate their initial reactions to events. Behavioral skills are targeted at increasing patient behaviors that can provide a sense of pleasure, accomplishment, and enhance problem-solving. Patients learn to engage proactively in pleasurable activities to boost their mood, break down problems into smaller parts, and to troubleshoot multiple ways of handling negative outcomes. These behavioral skills are intended to function in the service of providing patients with information and experience to target and question their dysfunctional attitudes. After treatment has concluded, highly skilled patients are hypothesized to maintain their gains by utilizing these skills when experiencing negative moods (Beck 1967, 1976). The acquisition of this skill set over the course of treatment has been hypothesized to be a

plausible mechanism whereby CT may achieve its enduring effects (Barber & DeRubeis, 1989).

Assessment of Patient Skills

Evidence for a relationship between CT skills and risk of relapse comes from a small set of studies that have largely relied on the Ways of Responding questionnaire (Barber & DeRubeis, 1992). The WOR was developed to measure patients' mastery of CT skills. The WOR asks participants to imagine six hypothetical scenarios and corresponding initial negative automatic thoughts. Using an open-ended response format, participants report what further thoughts they would have and what, if anything, they would do to cope. The first coder from a team of three raters initially parses responses into individual thought units. Using a detailed rating manual, the first two coders independently rate responses in terms of how much each statement reflects a positive (e.g., generating alternative explanations) or negative (e.g., placing blame on self) response, and the third rater resolves any disagreements. The resulting ratings are summarized through a total score, which is the number of positive category ratings given minus the number of negative category ratings. The three raters each provide a rating of overall response quality, reflecting how much each strategy would be expected to improve one's mood.

In an initial validation study, both the WOR Total and Quality scores were shown to correlate with lower levels of depressive symptoms among college students (Barber & DeRubeis, 1992). Improvement on the WOR over 12 weeks of treatment among depressed patients has been shown to coincide with concurrent reductions in depression symptoms (Barber & DeRubeis, 2001). In a separate study conducted among treatment responders, higher scores on the WOR were shown to predict risk for relapse at a 1-year follow-up even after controlling for post-treatment residual symptoms (Strunk et al., 2007).

Results using the WOR Positive and Negative scores have differed across two studies. In Barber and DeRubeis' sample of adult outpatients with MDD, improvement in WOR Positive scores over the course of treatment were associated with concurrent reductions in depressive symptoms (r = .57), but this relationship was not found for change in WOR Negative scores (2001). A subsequent study yielded an opposite pattern of results. Analyses using a pooled study database of adolescents and adults with either MDD, generalized anxiety disorder, panic disorder, borderline personality disorder, or adolescent anxiety disorders examined change in the WOR across cognitive and psychodynamic therapies and found only the WOR Negative scores to correlate with reductions in depressive symptoms (Connolly Gibbons et al., 2009). It is important to highlight that these discrepant results could be related to differences in the samples, as the pooled study database in Connolly Gibbons and colleagues' study comprised much greater heterogeneity in patient age, diagnosis, and treatment type than the sample used by Barber and DeRubeis (2001).

To complement the WOR based assessment of patients' mastery of CT skills, Strunk and colleagues (2007) used an observer-rated measure of patients' performance of these skills both in and between sessions – the Performance of Cognitive Therapy Skills scale (PCTS). Like the WOR, the PCTS also predicted risk of relapse in the year following treatment with CT. When entered in the same model, both the WOR and the PCTS were independent predictors of risk for relapse. While not reported in the main paper, Strunk and colleagues (2006) reported in a conference presentation that these two predictors, while uncorrelated (r = .11), served as a very strong predictor of risk for relapse when combined in a composite index.

Although available evidence suggests the WOR and the PCTS are useful predictors of risk of relapse, the lengthy rating process and multiple coders required to score these instruments make them labor-intensive for researchers and impractical for use in clinical practice. If a valid self-report measure of patients' mastery and use of CT skills could be developed, it would have obvious advantages in this regard. Moreover, in addition to predicting long-term outcomes, a short skills assessment could help therapists maximize early treatment outcomes by directing them to engage in more Socratic techniques based on patients' coping skills at the onset of treatment. With these goals in mind, we developed a client reported 30-item measure: the Patient Competencies Scale – Client version (PCS-C). We also developed a 9-item therapist rated measure: the Patient Competencies Scale – Therapist version (PCS-T). The language of the PCS-C was carefully constructed so that the measure could be used both pre and post-treatment. Items can be understood without any experience with CT or CT skills. Items on both measures strongly reflect cognitive and behavioral skill usage, with some coverage of patients' recognition of the inaccuracy of their core beliefs.

Skills of Cognitive Therapy Scale

In parallel with the development of our measure, Jarrett and colleagues (2011) have designed the Skills of Cognitive Therapy (SoCT) scale intended to evaluate patients' understanding and use of CT skills. This 8-item measure assesses CT skills from the perspective of both patients and observers. While questions focus primarily on a patient's ability to identify and reevaluate automatic negative thoughts, one item assesses the patient's understanding of core beliefs. In an initial study of acute phase CT among patients with recurrent MDD, therapist and patient SoCT scores from mid-point and post-treatment assessments were significantly associated with post treatment depressive symptoms and treatment response. However, because these analyses did not control for patients' initial depression severity, the extent to which these associations reflect either concurrent correlations of SoCT ratings and depressive symptoms or the association of change in both the SoCT rated skills and depressive symptoms remains unclear.

While the SoCT is designed to measure coping skills once they have been acquired in the course of CT, our measure was developed to evaluate coping skills both before a patient has begun treatment and after treatment is underway. While attention has been drawn to the utility of the assessment of CT skills at post-treatment (i.e., for the prediction of subsequent risk for relapse), there may also be clinical utility to the assessment of these CT skills prior to treatment. First, insofar as pre-treatment skill level may relate to severity of symptoms, these skills could indicate patients that may be more or less responsive to CT. In support of this notion, one study has found a relationship between pre-treatment skill usage and pre-treatment depressive symptoms. In a sample of college students assumed to be naïve to CT, use of CT skills predicted greater adaptive responses and less depressive symptom reactivity in response to repeated stressors (Adler, Conklin, & Strunk, under review). This finding suggests that skills comparable to those taught in CT are employed naturally by healthy individuals. Second, should there exist an interaction between greater pre-treatment CT skills and therapists' early use of Socratic questioning in predicting early symptom change, identifying patients with greater skills could afford therapists the potential to maximize outcomes for these patients. Because the SoCT is not intended for use pre-treatment, our PCS-C may have the advantage of better assessing changes in CT skills over the course of treatment and identifying patients who could benefit from more therapist use of Socratic questioning.

Furthermore, in addition to assessing competence in identifying and reevaluating automatic negative thoughts, our measure gives greater attention to evaluating individuals' modification of underlying core beliefs. The choice of Jarrett and colleagues to use only those items that they believed could be rated through observation of therapy sessions largely prohibited such an emphasis on core belief change (2011). The SoCT uses the same form to assess patient's CT skills from the perspectives of both observers and patients. In contrast, based on our reasoning that patients have access to different information about their thoughts, beliefs, and behaviors outside of therapy than their therapists, we developed a separate therapist and client version of the PCS with the expectation of capturing these distinct features of patients' competencies. Because the SoCT was introduced since the data for this study were collected, we were unable to use

both measures for comparison. However, we will return to the potential similarities and differences between the SoCT and PCS in the discussion.

Purpose

The present study has two primary goals. First, we will examine the validity of the client and therapist versions of the PCS in the hopes of developing new measures of CT skills. To do so, we will examine the factor structure of the client and therapist versions of the PCS. We will assess the extent that these PCS measures can detect changes in skill level from pre to post treatment that previous research using the WOR suggests occur. Next, we will investigate whether skill level differentiates patients with MDD from healthy controls both before patients with MDD have begun treatment and after completing a course of CT. We will evaluate the predictive validity of the PCS by comparing the correspondence of change on the PCS over the course of treatment with changes in complementary measures (e.g., WOR, DAS) as well as the ability of PCS changes to predict rates of symptom change and risk of relapse. Our second aim is to examine if there is an interaction between CT skills prior to beginning treatment and therapist use of Socratic techniques in predicting early symptom change. We hypothesize that patients with greater PCS rated skills at intake will benefit more in terms of early session-to-session symptom change from therapists' use of Socratic questioning.

Chapter 2: Method Participants

We examined two groups of participants: a sample of depressed patients who participated in 16 weeks of CT and a sample of control participants who were not depressed. Both groups were recruited for participation as part of a larger study (Adler & Strunk, personal communication). The depressed sample was comprised of 65 participants and was predominately female (55%) and Caucasian (85%). The average age of the depressed group was 36 years (SD = 13.3). Patients were included in the depressed sample who met the following inclusion criteria: (a) diagnosis of current MDD according to the Diagnostic and Statistical Manual for Mental Disorders (DSM-IV; American Psychiatric Association, 1994); (b) 18 years or older; (c) able and willing to give informed consent; and (d) agreed not to make changes to their psychiatric medication regimen over the course of the study. Individuals were excluded from the study if they had (e) a history of Bipolar I disorder or psychosis; (f) current, primary Axis I disorder besides MDD requiring treatment other than CT for depression; (g) substance dependence within the past six months; (h) IQ of 80 or below (with assessment only occurring when clinically indicated); (i) clear potential for secondary gain (i.e. court ordered treatment); (j) current suicide risk contraindicating treatment on an outpatient basis.

The control group was comprised of participants matched on sex and age (within 2.5 years; M = 38, SD = 14.2) and education to the 44 participants in the depressed sample who completed treatment. To participate as a control, participants needed to meet the following inclusion criteria: (a) no current or lifetime history of anxiety disorder,

mood disorder or psychosis per the Structured Clinical Interview for DSM-IV (SCID); (b) 18 years or older (in addition to matching the age of a participant in the depressed sample); and (c) able and willing to give informed consent. Potential participants were excluded from the study if: (d) they had an IQ of 80 or below (with assessment only occurring when clinically indicated).

Measures

Axis I Diagnosis

Structured Clinical Interview for DSM-IV (SCID). The SCID (First, Spitzer, Miriam, Williams, 2002) was used to assess if patients met formal criteria for current MDD and any other Axis I disorders.

Depressive Symptom Severity

Hamilton Rating Scale for Depression (HRSD). We used the HRSD (Hamilton, 1960; Williams 1988), a clinician-rated 17-item measure of depressive symptom severity modified to assess atypical symptoms. This continuous measure was used as our primary measure of depressive symptom severity. The HRSD has demonstrated adequate reliability as well as internal consistency ($\alpha \ge .70$; Bagby, Ryder, Schuller, & Marshall, 2004).

Beck Depression Inventory- 2nd Edition (BDI-II). The BDI-II (Beck, Steer, & Brown, 1996), a 21-item self-report instrument, was used as our secondary measure of depressive symptom severity. Respondents rate each item (e.g., pessimism, sadness, loss of pleasure) on a scale from 0 (not at all) to 3 (strongly) with respect to how they've been feeling in the past week. Possible scores range from 0 (minimal to no depression) to 63 (severe depression). This measure has repeatedly demonstrated excellent internal consistency, with coefficient alphas averaging .91 or higher (Dozois, Dobson, & Ahnberg, 1998; Steer, Kumar, Ranieri, & Beck, 1998; Sprinkle et al., 2002). Estimates of the convergent validity of the BDI-II have produced moderate to large correlations with alternative measures of depression severity (Beck et al., 1996; Dozois et al., 1998; Riskind et al., 1987).

Patient Competencies in CT Skills

Ways of Responding Inventory (WOR). The WOR (Barber & DeRubeis, 1992) measures the acquisition of a set of skills that cognitive therapists seek to foster in their patients. Respondents are provided with 6 hypothetical scenarios (e.g., being turned down for several jobs to which they applied) and corresponding initial thoughts (e.g., "There just doesn't seem to be any point in applying"). Respondents are asked to report how vividly they are imagining the situation, their mood intensity, and their ability to imagine having the initial thoughts. They are instructed to record any subsequent thoughts or behavioral reactions they would have about the situation and then are asked to re-rate the intensity of their mood. Using guidelines outlined in the WOR *Rater's Guide* (Barber & DeRubeis), three coders rate the subsequent thoughts and behaviors provided by respondents. Responses are parsed into individual thought units, and two independent raters classify each parsing as either a specific positive or negative response. A third rater judges any disagreement between the first two raters. A total score is calculated from these ratings by subtracting the number negative response from the number of positive responses. An additional summary score, quality, assesses on a 7point scale the extent that the response would improve an individual's mood. The quality score is calculated by averaging the scores of the three independent raters.

While the WOR has many desirable properties, it requires a team of trained raters and is therefore time-consuming to rate and score. In an effort to assess patient competencies in CT in a manner that would allow for quick scoring, we developed the Patient Competencies Scales.

Patient Competencies Scales (PCS). The client version of the *Patient Competencies Scale* (PCS-C) is a 30 item self-report instrument. The therapist version of the *Patient Competencies Scale* (PCS-T) is somewhat shorter with 9 items. Both measures are designed to assess the clients' mastery of the specific skills that therapists seek to help clients develop in CT. Instructions ask respondents to indicate what would be true in the past two weeks. Each item is rated on a 0 (not at all) to 6 (completely) scale. The items of the PCS-C assess client's usage of behavioral skills, including engagement in pleasurable activities and problem solving practices, as well as cognitive skills including understanding the connection between thoughts and mood, and identifying, evaluating, and reassessing automatic thoughts. Items of the PCS-T evaluate a client's ability, independence, and frequency of use of behavioral activation, automatic thoughts, and core belief related strategies.

Dysfunctional Attitudes Scale (DAS). The DAS (Weissman, 1979) is a 40-item self-report instrument measuring maladaptive beliefs ("I cannot be happy unless most people I know admire me") common among depressed individuals as proposed in Beck's theory. While this is not considered a direct measure of effortful CT skill usage, the DAS

offers a more direct assessment of patients' beliefs and underlying assumptions that is likely to relate to their competency in CT skills. Correlating the DAS with the BDI-II has provided strong support for its convergent validity (r = .41; Cane, Olinger, Gotlib, & Kuiper, 1985; Oliver & Baumgart, 1985). Internal consistency ($\alpha = .87$) has been shown to be adequate.

Depression Relapse

Longitudinal Interview Follow-Up Evaluation (LIFE). The LIFE (Keller et al., 1987) is a clinician rated inventory used to determine depression relapse. Depression severity is rated on a 6-point scale for each week preceding the administration of the LIFE. A patient was determined to have relapsed if they met DSM-IV criteria for a Major Depressive Episode (as rated by a score of 5 for 2 consecutive weeks and requiring the endorsement of either loss of pleasure or depressed mood) at any point after terminating treatment. Date of relapse was estimated using the LIFE, and time to relapse after completing treatment (rounded to the nearest quarter month) was used in analyses. The interview was administered to treatment responders at 6 and 12 month follow-ups after the termination of treatment.

Socratic Questioning

Socratic Questioning Scale (SQS). The SQS (Braun & Strunk, personal communication) is comprised of 8 Likert-style questions on a 7 point-scale, with higher scores reflecting greater therapist implementation of Socratic techniques. Trained raters use the scale to assess therapist's adherence to systematic questioning and disavowal of knowledge, both process variables theorized to be related to the Socratic method

(Overholser, 1991; 1993a; 1993b; 1994; 1995). Systematic questioning is a term given to the therapist's use of a series of questions to guide the client through a process of selfdiscovery and foster client autonomy (Overholser, 1993a), and disavowal of knowledge refers to the therapist's efforts to encourage the client to view their beliefs as subjective experiences rather than objective facts (1995).

Procedures

Assessment measures were given at multiple time points throughout the study. Interview-based measures (SCID, HRSD) and self-report questionnaires (BDI-II, WOR, DAS, PCS-C) were administered at the initial evaluation. Participants completed a BDI-II before each therapy session, and the HRSD and DAS were given at the fourth week of treatment. SQS ratings were conducted for the first three sessions. At the post-treatment assessment (which occurred 16 weeks after the intake assessment), interviewers administered the HRSD, several self-report measures (BDI-II, WOR, DAS), and portions of the SCID assessing disorders for which each participant had met diagnostic criteria during the initial evaluation. The LIFE was administered during 6 and 12 month followup evaluations after treatment termination. Control participants completed the DAS, WOR, PCS-C, BDI-II, HRSD, & SCID during their initial assessment.

Chapter 3: Results Validation analyses for the PCS-C and PCS-T

The Factor Structure of the PCS

Our first primary aim was to examine the validity of our measures of patient competencies in CT skills, one version to be rated by clients (PCS-C) and a second version to be rated by therapists (PCS-T). To do so, we first conducted an exploratory factor analysis to examine the factor structure of the PCS-C. In keeping with recommendations from Fabrigar and colleagues (1999), we planned to conduct a principal factor analysis with Promax rotation. On the basis of the scree plot and parallel analysis, the factor analysis of the 30 items on the PCS-C identified a one factor solution. All items had factor loadings of .4 or above except item 15 (factor loading of .096). Due to the low factor loading of this item, we removed item 15 and repeated the Exploratory Factor Analysis (EFA) with the remaining 29 items. Initial inspection of the scree plot suggested a one factor solution. To investigate further the appropriate number of factors, we conducted a parallel analysis which creates a random data set using the same number of variables and observations as our data set. After computing the correlation matrix and corresponding eigenvalues for this random dataset, we compared the eigenvalues from our EFA to the eigenvalues from this random dataset. When eigenvalues from the random dataset exceed the eigenvalues from our observed data, it suggests that the factors are largely noise and should not be retained. The parallel analysis based on 29 items also identified a one factor solution. The first two eigenvalue estimates from the parallel analysis were 2.6 and 2.31 (assuming 77 observations and 29 random variables). The first two eigenvalues from the factor analysis were 9.47 and 1.97. Therefore, the

eigenvalue for the first factor of the factor analysis exceeded only the corresponding first factor of the parallel analysis, again suggesting a one factor solution for the PCS-C.

We then conducted an EFA of the 9 items of the PCS-T. Examination of both the scree plot and parallel analysis suggested a one factor solution. The first two eigenvalue estimates in the parallel analysis (assuming 27 observations and 10 random variables) were 1.96 and 1.61. Only the eigenvalue for the first factor (6.86) of the factor analysis but not the second (.73) exceeded the corresponding estimates from this parallel analysis, indicating a one factor solution. Because all items had a factor loading of .82 or above, all 9 items were retained. Scores on the PCS-C post treatment were significantly correlated with PCS-T scores at post treatment (r = .54, n = 40, p = .0003). Finding both the PCS-C and PCS-T are assessing one factor suggests that each scale is measuring a single underlying construct, ostensibly, patients' CT skills. While both scales were designed to measure three main aspects of CT skills (patients' competency in cognitive, behavioral, and core belief change strategies), a one factor solution indicates that these skill sets are all capturing the same underlying construct.

Detecting Change from Pre to Post Treatment

Next, we sought to determine if patients experienced a significant change in CT skills as assessed by the PCS-C between intake and post treatment. A one sample t-test analysis assessing differences between PCS-C scores at intake and post treatment indicated a large, significant change (d = 1.41) in scores between the two time points (t (41) = 6.10, p < .0001). This suggests that the comparably shorter PCS-C is capable of detecting

changes in coping skills over the course of CT treatment that have been found previously using lengthier measures (the WOR) of skills.

Differentiating between Patients with MDD and Controls

We also compared the PCS-C scores of matched controls with patients' scores both at intake and post treatment using paired t-tests to examine if PCS-C scores differed across these groups. We found a statistically significant difference between control PCS-C scores and patient scores at intake (t (120) = 5.82, p < .0001). The effect size for this difference was large, having a Cohen's d of 1.06. At post treatment, the difference between controls' scores and patients' PCS-C scores was small (d = ..12) and no longer significant (t (86) = ..55, p = ..58). Therefore, the PCS-C distinguished between patients with MDD and never-depressed controls in terms of their skill level prior to undergoing treatment. Because these differences in skill were no longer detectable using the PCS-C after patients with MDD had completed the course of acute CT, this finding suggests that MDD patients who complete 16 weeks of CT are similar to non-depressed individuals with regard to their performance and mastery of coping skills assessed by the PCS-C.

Correlations with Depression Severity and Related Skills Measures

We examined the concurrent associations of PCS scores and conceptually related measures of CT skills and depression severity at both intake and post treatment evaluation. These analyses examined the relationship between scores on the PCS-C with the WOR, the DAS, the BDI-II, and the HRSD at both intake and post treatment assessments (see Table 1). Because the PCS-T was administered only at post treatment, correlations between the PCS-T and these measures were limited to this assessment point (see Table 1). The PCS-C was related to both measures of depression symptom severity at intake (only at trend level for the HRSD) and post treatment, and significant associations were found between the PCS-T and both measures of depression severity at post treatment. These associations suggest that PCS rated skill level parallels patients' degree of depression symptom severity both when beginning and on completing treatment.

PCS-C scores at both intake and post treatment showed moderate, significant correlations with WOR Total and Quality ratings at concurrent time points. Both WOR Quality and Total scores were also related to the PCS-T at the post-treatment assessment. The correlations between both the therapist and client versions of the PCS with the WOR suggest that these three measures are assessing some overlapping aspects of patients' CT skills. However, the moderate size of these relationships indicates that the PCS and WOR might also also assess distinct features of CT skills. This inference is not unexpected given that the WOR instructs patients to respond to hypothetical stressful events and negative thoughts, while the PCS-C asks patients to report their actual performance of these coping skills and the PCS-T instructs therapists to assess patients' frequency, competency, and independent use of these skills.

To explore further the relationship between both the PCS-C and the PCS-T with the WOR scores, we examined relations with the Negative and Positive ratings on the WOR. PCS-C scores at intake (r = -.38, n = 63, p = .002) and post-treatment (r = -.37, n = 43, p = .02) showed moderate, significant associations with WOR Negative scores at concurrent times of assessment, but these relationships were smaller and non-significant when examining WOR Positive scores at both time points. The PCS-T at post-treatment showed a moderate, trend relationship with the WOR Negative (r = -.29, n = 42, p = .06) but not the Positive scores. Therefore, it appears that the relationship between the PCS and the WOR is in part a reflection of the extent to which these measures capture variability associated with patients' experience of negative cognitions in response to stressors.

When examining the correlation between the PCS and a measure of dysfunctional attitudes, the PCS-C was not significantly related to the DAS at intake, but DAS scores post treatment showed moderate, significant associations with PCS-C post scores and were associated at trend level with PCS-T post scores. Thus, PCS-C rated skills are unrelated to patients' endorsement of negative, maladaptive beliefs prior to beginning treatment, but patient ratings of these beliefs show a modest relationship with therapist and client rated skills on the PCS after completing treatment.

In an investigation of the predictive validity of our PCS measures, we examined whether the acquisition of these patient reported competencies corresponded to changes on the WOR and DAS that previous research has demonstrated to occur over the course of treatment. To do so, we calculated residualized change scores (i.e., residuals from regression models where the intake score on one measure predicted the post score on the same measure) with the WOR and the DAS. We calculated the correlations between residualized change scores on the WOR and DAS and both the PCS-C and PCS-T residualized change scores. Because therapists' ratings of patient skills at intake were not assessed, residualized change scores for the PCS-T were calculated using the PCS-C scores at intake.

Findings from these analyses are presented in Table 2. Change in PCS-C and PCS-T showed a moderate, significant relationship with WOR Quality, and while the association with WOR Total was also moderate, this was a non-significant trend. This first finding indicates that skill acquisition measured by the PCS-C and PCS-T was related to the reporting a set of coping responses on the WOR which raters viewed as likely to improve their mood. When comparing the association between WOR Positive and Negative scores with PCS residualized change, we found the same pattern of results as when we examined PCS scores at intake and post treatment. The PCS-C and PCS-T change scores were moderately correlated with the sum of negative responses on the WOR, but, as with PCS intake and post ratings, this association was not found with the WOR Positive score, indicating that the number of patient responses coded as positive were unrelated to residualized change on either PCS measure.

PCS-C and PCS-T residualized change scores showed a significant moderate association with the DAS. In summary, change in patient skill over treatment as assessed by residualized scores on both the PCS-C and PCS-T paralleled change in the DAS scores as well as change in some but not all of the WOR ratings.

Predicting Change in Depression Severity

Using the two measures of depression severity available at intake, week 4, and post treatment, two hierarchical linear models (HLMs) were used to examine residualized change on the PCS-C and PCS-T separately as predictors of change in the BDI-II or

HRSD over the course of treatment (see Table 2). There was evidence for nonlinearity for the rate of symptom change on the BDI-II, so we examined several variations of the time variable including the log transformation and square root transformation of time. A comparison of fit statistics indicated that models involving the square root transformation of time demonstrated the best model fit, so we used this transformation for analyses. The standard linear representation of time was used for models with the HRSD because there was no evidence of nonlinearity, and fit statistics indicated this provided the best model fit. The interaction of residualized change on the PCS-C by Time was a large, significant predictor of change in both the BDI-II and HRSD, indicating that greater CT skill acquisition was related to a faster rate of change in depressive symptoms (See Table 2). The interaction of residualized change on the PCS-T by Time was also a moderate, significant predictor of change on the BDI-II but predicted at a non-significant trend for the HRSD.

After examining whether change on the PCS-C predicted patients' slopes of symptom change over the course of treatment, we examined three separate models to investigate the incremental validity of the PCS-C and PCS-T relative to the WOR in predicting change in depression severity across treatment. Model 1 included three predictors, residualized change scores for the PCS-C, PCS-T, and WOR Quality ratings, as well interactions of these predictors with time. Although our previous models examining the main effects of residualized change scores for the PCS-C and PCS-C and PCS-T (see Table 2) separately found both measures to be independent predictors of depressive symptom change on the BDI-II and HRSD, Model 1 (see Table 3) with our three

predictors found only the interaction of PCS-C by Time to be a unique, significant predictor of change on the HRSD (t (37) = -2.30, r = .35, p = .03) and a trend level predictor of change on the BDI-II (t (26) = -1.98, r = .36, p = .06). We included residualized change scores for WOR Total instead of WOR Quality ratings and conducted further analyses with this modified version of Model 1 (See Table 4). Though results using this modified Model 1 were comparable, some discrepancies emerged when incorporating the alternative scoring system for the WOR into the model. When predicting change in HRSD over time, both the interaction of the PCS-C and Time (t (37) = -2.39, r = .37, p = .02) and the interaction of WOR Total and Time (t (37) = -2.02, r = .32, p = .05) emerged as significant predictors of change in HRSD scores over the course of treatment. A corresponding model (see Table 4) predicting slopes of change in BDI-II scores using the WOR Total scores in place of the WOR Quality scores found the interaction of PCS-C scores with Time to be predictive of BDI-II change at trend (t (26) = -1.94, r = .36, p = .06), and none of the other predictors reached statistical significance.

After examining our three predictors of interest in Model 1 in combination, we sought to compare the unique predictive ability of the WOR in predicting symptom change against the PCS-T and PCS-C separately in Model 2 and 3 respectively. Model 2 was comprised of residualized change on the PCS-T and WOR Quality along with the interactions between each of these predictors and Time (see Table 3). In the model predicting change in the BDI-II, WOR Quality was not statistically significant (t (26) = -.33, r = .10, p = .62,), and the interaction of residualized change on the PCS-T by Time reached only trend level significance (t (26) = -1.79, r = .33, p = .08). The same model

predicting change in HRSD found none of our predictors to be significant predictors of symptom change. A modified version of Model 2 including WOR Total ratings in lieu of WOR Quality ratings (see Table 4) found a comparable pattern of results in predicting change in the BDI-II. While none of the predictors were statistically significant, the interaction of residualized change on the PCS-T and Time reached trend level significance (t (26) = -1.78, r = .33, p = .09,). However, in an equivalently modified model predicting change in the HRSD, the interaction between PCS-T and Time was non-significant (t (29) = -1.23, r = .19, p = .23,), but the interaction between WOR Total and Time was a significant predictor of HRSD change (t (39) = -2.16, r = .33, p = .04,). This pattern of results indicates that, after accounting for the variability in patients' CT skills that the PCS-T and WOR assess in common, the remaining variability in PCS-T ratings did not predict change in depressive symptoms over the course of treatment.

Finally, we investigated a model to examine the unique predictive power of the WOR against the PCS-C in predicting change in depressive symptom severity. Model 3 included residualized change on the PCS-C and WOR Quality as well as interactions between each of these two predictors and Time (see Table 3). The interaction of residualized change on the PCS-C by Time was a significant predictor of change in BDI-II scores (t (27) = -2.52, r = .44, p = .02,), whereas the interactions of the WOR Quality score by Time was not statistically significant (t(27) = -.23, r = .04, p = .82). When predicting change in the HRSD, we found a comparable pattern of results such that the interaction of PCS-C by Time (t(40) = -2.66, r = .39, p = .01) was significant while the interaction of WOR Quality by Time was not (t(40) = -1.35, r = .21, p = .82). A modified

version of this model using WOR Total scores found the same pattern of results (see Table 4), with only the interaction of PCS-C by Time reaching significance in predicting change in BDI-II scores (t(27) = -2.46, r = .43, p = .02). An equivalent model predicting change in HRSD scores found the interaction of residualized PCS-C change scores by time to be a significant predictor of symptom reduction on the HRSD across treatment (t(40) = -2.77, r = .40, p = .01), while the residualized WOR Quality scores by Time interaction reached statistical significance at trend level (t(40) = -1.85, r = .28, p = .07). Therefore, in contrast to the PCS-T which did not predict change in depressive symptoms above and beyond the WOR, the client rated PCS scores emerged as a consistent predictor of symptom reductions across treatment after accounting for change in the WOR rated skills. This finding signifies that client ratings of their coping skills on the PCS-C capture variability predictive of symptom reduction above and beyond the aspects of CT skills captured by patient responses on the WOR, but therapist ratings on the PCS-T are not uniquely predictive of change in depression symptoms after considering the variability in symptom change accounted for by the WOR.

Predicting Risk of Relapse

We conducted survival analyses to examine the relationship between our PCS measures at post treatment with patients' risk of relapse. Twenty-nine patients were identified as having responded to the treatment, and of those that responded, 12 were identified as having relapsed. Six patients were censored prior to the end of the 12 month follow up assessment (due to either being lost to follow-up or seeking additional treatment despite not having met relapse criteria). We identified a list of 6 potential

covariates likely to be related to relapse (BDI-II and HRSD at post and residualized change controlling for intake severity for both of these measures, recurrent MDD, and patients' age) to include in our model and predetermined to retain any with p-values less than .1. We found age and HRSD at post treatment to be significant predictors of relapse, so these covariates were included in our model with skills as a predictor of relapse. Survival analyses including these covariates indicated that neither the client PCS-C scores at post treatment, $\chi^2(1, n = 28) = .0008$, hazard = 1.00, 95% Confidence Interval (CI) = [.98, 1.02], p = .98, nor the therapist PCS-T post treatment scores, $\chi^2(1, n = 27) =$.36, hazard = 1.21, 95% CI = [.65, 2.27], p = .55, predicted relapse. When eliminating age and HRSD post treatment scores as covariates, separate models using PCS-C and PCS-T scores as predictors found neither to predict relapse (*ps* greater than .4). Thus, while the PCS-C and PCS-T both independently predicted reductions in depressive symptoms over the course of treatment, these measures failed to capture variability in skills that related to patients risk for relapse in the 12 months following treatment termination.

Socratic Questioning Analyses

For our second primary goal, we next turned our attention to examining the relationship between the PCS-C ratings at intake and therapists' early use of Socratic questioning in predicting patients' session to session symptom change over the first four sessions of CT. We conducted repeated measures regression analyses to investigate the relationship between Socratic questioning and PCS-C scores at intake in predicting session-to-session symptom change as measured by the BDI-II over the first three

sessions of treatment. Using SAS Proc Mixed, we used data for each of the first three sessions in a single omnibus test to analyze the relationship between SQS ratings and symptom change in the following session as well as the interaction between PCS-C scores at intake and these SQS ratings. This procedure includes SQS scores for each session (1-3) as predictors of BDI-II scores in the immediately following session (1-3). We included BDI-II scores from the current session, PCS-C scores at intake, a dichotomous variable indicating whether or not the patient was taking ADM, and the interaction of the session's SQS score and intake PCS-C scores as covariates in these analyses. Thus, a vector of BDI-II scores from sessions 1 through 3 served as the dependent variable, and a vector of BDI-II scores from the prior session (1-2) served as one of the covariates, allowing the BDI-II scores from a prior session to be covaried out of the BDI-II ratings at the following session. Four covariance structures were compared using AIC (autoregressive, compound symmetry, toeplitz, unstructured), and unstructured was identified as the model with the best fit.

The interaction between PCS-C at intake and the previous session SQS scores reached significance at trend in predicting session-to-session symptom change over the first three sessions (t (148) = -1.71, r = .14, p = .09). To investigate further the nature of this interaction, we identified patients as having either high or low baseline skills based on a median split of PCS-C scores and examined the effect of Socratic questioning on symptom change for these high and low PCS-C groups separately. Among those with lower baseline skills on the PCS-C, Socratic questioning was not significantly related to greater symptom reduction (t (76) = -.36, r = .04, p = .72). For patients with higher

baseline skills, greater use of Socratic questioning was significantly related to greater decreases in depressive symptoms (t(70) = -2.28, r = .23, p = .03). This pattern of results raises the possibility that patients with greater baseline skills might benefit more in terms of early session symptom reduction from therapists' greater use of Socratic techniques early in treatment.

Chapter 4: Discussion

This study was conducted in an effort to evaluate a newly developed self-report and therapist rated assessment of CT specific skills targeted for MDD patients to acquire during the course of CT for depression as well as to examine if there was an interaction between therapists' use of Socratic questioning and client rated CT skills at the onset of therapy in predicting early session symptom improvement. EFAs conducted separately for the PCS-C and PCS-T each yielded one factor solutions, indicating that these scales each assess a single underlying construct despite comprising questions specific to cognitive, behavioral, and core belief skills. A previous study evaluating another brief, self-report measure of patients' CT skills, the SoCT, also reported a one factor solution for their scale using a substantially larger sample (N > 523; Jarrett et al., 2011). Both the PCS-C and PCS-T demonstrated good construct validity, showing moderate correlations with related measures of CT skills (WOR) and dysfunctional attitudes (DAS) as well as large correlations with measures of depressive symptoms (HRSD & BDI-II) assessed at comparable time points. Moreover, change in skills on the PCS-C and PCS-T was also related to change in related skill measures (WOR and DAS) as well as depressive symptom change as measured by the BDI-II. Change in the PCS-C showed a moderate, significant relationship with change in the HRSD, but the relationship between PCS-T and HRSD change reached only trend level significance. These moderately sized correlations suggest that the PCS-C and PCS-T are assessing some features of CT skills and depression severity similarly measured by these related scales, but our scales might also be measuring aspects of CT skills that are not assessed by these related measures. This finding is not unexpected given that our measures were designed to assess patients'

competency and independent use of these coping skills outside of therapy, which was likely not an explicit aim of the authors of these other measures.

In addition, while higher skills as assessed on both measures were related to greater use of negative coping strategies on the WOR Negative scale, neither scale was associated with ratings of positive coping strategies on the WOR Positive scale. The WOR negative scale assesses patients' use of negatively distorted and often self-blaming automatic thoughts in response to hypothetical stressors, whereas the WOR positive scale reflects patients' ability to examine evidence and generate alternative responses to these situations. Finding an association between the PCS measures and the WOR negative but not positive scale could indicate that, as patients' depressive symptoms decrease, these individuals experience fewer of these self-critical thoughts but may not be coping by actively using the cognitive restructuring techniques taught in CT. In support of this interpretation, we conducted additional exploratory analyses not reported above and found that the association between WOR Quality ratings and both PCS measures (for PCS-C: r = .37, p = .01; for PCS-T: r = .35, p = .02) did not remain significant after controlling for WOR negative scores (for PCS-C: r = .17, p = .26; for PCS-T: r = .21, p =.18). This suggests that the relationship between WOR Quality scores, which are intended to indicate the likelihood that a response would improve the patient's mood, may be the result of experiencing fewer negative automatic thoughts rather than restructuring these cognitions in the face of stressors.

As would be predicted by a theoretical framework proposing skill acquisition to be related to symptom reduction in MDD, patients' PCS-C scores showed statistically significant improvement from intake to completing treatment. This finding demonstrates that the PCS-C detected changes in CT skills that occur over the course of CT. Furthermore, in support of the construct validity of the PCS-C, this measure discriminated between MDD patients and non-depressed matched controls in level of skills at intake. At post treatment, no significant differences were found between the MDD patients and controls. These findings suggest that patients in CT are developing a set of skills that healthy individuals employ naturally.

A novel finding of this study is the interaction demonstrated between pretreatment skill levels and therapist use of Socratic questioning in predicting early session symptom change. We found that patients with greater pre-treatment competency with these skills benefited more in terms of symptom reduction from the use of Socratic questioning techniques than did patients with comparatively poorer skills usage. Such an interaction suggests that therapists' proficient usage of these techniques may have the potential to maximize outcomes for those patients with greater skills prior to beginning treatment.

We examined three models to assess the incremental validity of our PCS measures against the WOR in predicting change in depressive symptoms. In the first model, we included the WOR, PCS-C, and PCS-T together as predictors (along with their interactions with Time). Using this combined model, only the PCS-C emerged as a consistent predictor of change in both measures of depressive symptom severity (BDI-II and HRSD). Thus, the PCS-C was capturing variance related to symptom reduction in excess of that accounted for by the WOR assessment of CT skills. In our next models,

we examined the unique predictive power of the WOR against the PCS-C and PCS-T separately. Analyses with these models found that the PCS-C was significantly predictive of symptom change above and beyond the WOR across depressive symptom measures, whereas the PCS-T reached only trend level significance. In summary, across these three models, only the PCS-C consistently emerged as a significant predictor of short-term change in depressive symptoms in combination with other predictors. Over the course of treatment, the PCS-C accounted for significant variance in change in HRSD and BDI-II scores above and beyond that accounted for by WOR scores. While the WOR and PCS-C are conceptually related and moderately correlated, the PCS-C captures features of patients' CT skills that are both absent from the WOR and predictive of changes in depressive symptom severity. This finding supports the use of the PCS-C for the prediction of patients' symptom reduction of the course of therapy.

Although the PCS-C predicted symptom change, it was not predictive of risk of relapse in our sample. Survival analyses revealed that neither the therapist nor client version of our skills measure predicted likelihood of relapse over the 12-month follow-up period, whereas the WOR was found to be a significant predictor of risk of relapse in our sample. We posit three possible interpretations to explain this pattern of findings. First, this finding could indicate that the PCS does not capture variance associated with patients' skill level that is related to relapse and therefore is not an appropriate index for relapse risk. Alternatively, the limited number of patients who completed and responded to CT (n = 29) may have left the study underpowered to detect a relationship between PCS scores and relapse. While the PCS scores were not a significant predictor of relapse,

hazard ratios for our study indicate a small effect size. In contrast, in a comparably sized sample, Strunk and colleagues (2007) found coping skills (as assessed by the WOR) to be predictive of risk of relapse and reported a large effect size for this relationship. The discrepancy in the size of the effect across these studies suggests that our failure to find an effect is unlikely to have resulted from a lack of statistical power. One possible explanation of these discrepant findings is that our samples may have differed in patient characteristics that are meaningfully related to relapse risk. For example, our sample was comprised of patients who met criteria for MDD whereas the research by Strunk and colleagues' (2007) included only patients identified as having moderate to severe MDD. Should patients' initial severity of depressive symptoms be meaningfully related to risk of relapse, this difference could have the potential to have affected the relationship observed between patients' CT skills and risk for relapse in the two studies. Although we were unable to detect an effect of intake depression severity on the predictive ability of PCS rated skills in the current study, our power to detect such an effect in our current sample was low.

In summary, these findings are suggestive of the utility of the PCS-C and PCS-T as short alternatives to the WOR for measuring patient competencies and predicting short-term reductions in depression severity over the course of treatment. While, these measures have not demonstrated the ability to identify patients at risk for relapse, they offer a valuable, clinically useful method for assessing patient competencies in CT skills that have been related to change in depressive symptoms over the course of treatment. These shorter term reductions in depression severity remain informative outcomes for therapists to consider.

Nonetheless, given the importance of identifying patients at risk of relapse, we will be looking to alternative methods of assessing skill acquisition and usage. It seems likely, after undergoing a course of CT, a majority of patients will be capable of identifying numerous CT skills but may lack the ability to reproduce these skill sets properly or consistently when needed. We suspect that a structured clinician rated interview for assessing a patient's skills may be necessary to distinguish between the passive endorsements of skill usage and demonstrated competency in the face of real life stressors. Such an approach could reduce the influence of patient response biases and demand characteristics associated with self-report formats. In addition, therapist judgment may be required in cases when patients lack good comparisons of skill usage.

In an early attempt to assess patients' independent understanding and performance of CT competencies, Strunk and colleagues (2007) developed the PCTS. The PCTS relies on the assumption that trained raters can use patients' in session behaviors, including their reports of past behaviors and stated intentions for future behaviors, to predict their use of CT skills outside of therapy. While both these PCTS ratings and the WOR total score were found to be predictive of risk of relapse in the year following treatment, these two skill measures were only minimally correlated (r = .11; Strunk et al., 2007). Although our finding that the PCS-T failed to predict relapse could be taken to suggest that clinicians may not have a good sense of their patients' skill level, this finding using the PCTS leads us to suspect that a structured interview would give clinicians better access to the relevant information to draw these judgments. While clinicians spend substantial time helping patients develop these skills, they may spend comparatively less effort examining their patients' frequency and independent skill use or their ability to apply these skills across situations and in the event of stressors. As probability of relapse is not traditionally the primary purpose of a therapist's assessment, we suspect that an ad hoc measure such as the PCS-T may not be as useful in predicting this variable as a standardized clinician rated interview with targeted questions for eliciting evidence of demonstrated skill usage. Our hope is that a clinician rated measure that is slightly less labor intensive than the PCTS but more structured than the PCS-T may be both clinically useful and capable of predicting relapse.

It is important to note that the PCS-C is only one of multiple measures recently developed with the intent of measuring patients' competency in CT specific skills. The assessment most comparable to our PCS is the Skills of Cognitive Thearpy (SoCT; Jarrett et al. 2011). Although we lack a sample on which we can directly compare these measures, we will discuss briefly the relevant empirical support and clinical utility of this measure of skills. Jarrett et al.'s (2011) SoCT has been validated in a sample of depressed patients and was found to predict response to CT. While this 8-question measure focuses heavily on recognizing the connection between thoughts, feelings, and behaviors, comparably little attention is given to assessing a patient's skillfulness at identifying and correcting negative automatic thoughts, the quality of their behavioral skills, or to their ability to restructure maladaptive core beliefs. Alternatively, our PCS-C includes questions targeted at differentiating the quality of skill usage as well as the

presence of higher-level schema restructuring. In addition, unlike the PCS-C which can be rated by patients both before and after receiving treatment, the wording of the SoCT makes it unsuitable for administering to patients naïve to CT. Consequently, clinicians cannot assess patients' change in SoCT scores, whereas such assessments of skill change are possible using the PCS-C. In conclusion, while neither the PCS-C nor the SoCT have demonstrated the ability to predict relapse, we believe that the PCS-C is a more comprehensive inventory of a patient's competency in CT due to its capability of assessing change in skill level from pre to post treatment and its item content that affords greater focus to CT skills related to cognitive strategies and schema change.

Limitations

This study was subject to several limitations that should be noted. As mentioned previously, the PCS-C and PCS-T were only administered to a limited sized sample of patients, their respective therapist, and matched non-depressed controls. Furthermore, the format of the PCS-T made it only suitable for administration at post treatment, so we were unable to assess therapist rated changes in skills from pre to post treatment using this measure. Finally, changes in CT skills on the PCS-C were measured concurrently with related measures of skills (WOR and DAS) and depression symptom severity ratings (BDI-II and HRSD) rather than manipulated experimentally. While these findings provide compelling evidence that the development of skills shown on the PCS-C and PCS-T parallel changes detected on numerous relevant assessments, our data do not allow us to infer the potential causal nature of such a relationship. The one variable we had the capability of predicting, MDD relapse, was not found to be predicted by either

PCS-C or PCS-T scores, so we are unable to conclude whether our measures lack the ability to predict this outcome or if our data failed to identify a relationship between skills and relapse.

In conclusion, this study has begun the process of validating two measures of patient's skill use, a self-report and clinician-rated version, and our findings suggest that these measures could prove useful in both research and clinical settings as short, easy to use alternatives to existing measures. Both measures were found to be associated with related measures of skills, and skill acquisition paralleled improvement in depressive symptoms over the first several weeks of therapy. Skill level differentiated patients from healthy controls at pre-but not post treatment. Furthermore, we found pre-treatment measures of skills on the self-report version to interact with therapist use of Socratic questioning in the prediction of symptom change early in the course of treatment. Patients with greater skills at the onset of treatment benefited more in terms of depressive symptom reduction from the Socratic Method. This finding is important insofar as it suggests a way that patients' pre-treatment coping skills may impact early outcomes in CT. Moreover, this finding highlights a potential method whereby differential implementation of CT through greater use of Socratic techniques may lead to disparate patient outcomes and produce variability in outcomes for CT. Finally, neither measure was predictive of relapse. Although previous research has reported CT skills as measured by the WOR and the PCTS to be predictive of relapse (Strunk et al., 2007), neither of our self-report measures of patient competencies were related to risk of relapse during the 12 month follow-up period. Finding no relation between our measures and relapse could

indicate either that these measures are capable of detecting such a relationship or that no true relationship exists between skills and relapse for our measures to detect.Nonetheless, our hope is that a structured clinician-rated interview, which is currently being piloted, might be capable of capturing the variability in skills necessary to predict this long-term outcome.

References

- American Psychiatric Association. (1994). *Diagnostic and statistical manual of mental disorders*, Fourth Edition. Washington, DC: American Psychiatric Association.
- Adler, A.D., Conklin, L. & Strunk, D.R. (under review). Naturalistic use of Cognitive Therapy skills: Predicting depressive symptom reactivity

Adler, A. D., & Strunk, D. R. (personal communication).

- Bagby, R.M., Ryder, A.G., Schuller, D.R., & Marshall, M.B. (2004). The Hamilton Depression Rating Scale: Has the gold standard become a lead weight? *The American Journal of Psychiatry*, 161, 2163-2177.
- Barber, J. P., & DeRubeis, R. J. (1989). On second thought: Where the action is in cognitive therapy for depression. *Cognitive Therapy and Research*, 13, 441-457.
- Barber, J. P., & DeRubeis, R. J. (1992). The ways of responding: A scale to assess compensatory skills taught in cognitive therapy. *Behavioral Assessment*, 14, 93-115.
- Barber, J. P., & DeRubeis, R. J. (2001). Change in compensatory skills in cognitive therapy for depression. *Journal of Psychotherapy Practice & Research*, 10, 8-13.
- Beck, A. T. (1967). Depression: Clinical, experimental, and theoretical aspects. New York: Hoeber. Republished as Depression: Causes and treatment. Philadelphia: University of Pennsylvania Press.
- Beck, A. T. (1976). *Cognitive therapy and emotional disorders*. New York: International Universities Press.
- Beck, A.T., Rush, A.J., Shaw, B.F., & Emery, G. (1979) Cognitive Therapy of Depression. New York, NY: Guilford.

- Beck, A. T., Steer, R. A., & Brown, B. K. (1996). Beck depression inventory manual (2nd edition). San Antonio, TX: Psychological Corporation.
- Braun, J.D. & Strunk, D.R. (personal communication). Socratic method and therapist adherence as predictors of symptom change in Cognitive Therapy for depression: A study of therapists in training (*working title*).
- Cane, D. B., Olinger, L. J., Gotlib, I. H., & Kuiper, N. A. (1986). Factor structure of the Dysfunctional Attitude Scale in a student population. *Journal of Clinical Psychology*, 42, 307-309.
- Connolly Gibbons, M. B., Crits-Christoph, P., Barber, J. P., Wiltsey Stirman, S.,Gallop,
 R., Goldstein, L. A., Temes, C.M., & Ring-Kurtz, S. (2009). Unique and common mechanisms of change across cognitive and dynamic psychotherapies. *Journal of Consulting and Clinical Psychology*, 77, 801-813.
- DeRubeis, R. J., Gelfand, L. A., Tang, T. Z., & Simons, A. D. (1999). Medications versus cognitive behavior therapy for severely depressed outpatients: Mega-analysis of four randomized comparisons. *The American Journal of Psychiatry*, 156(7), 1007-1013.
- DeRubeis, R. J., Hollon, S. D., Amsterdam, J. D., Shelton, R. C., Young, P. R., Salomon,
 R.M. et al. (2005). Cognitive Therapy vs. medications in the treatment of
 moderate to severe depression. *Archives of General Psychiatry*, 62(4), 409-416.
- Dimidjian, S. et al. (2006). Randomized trial of behavioral activation, cognitive therapy, and antidepressant medication in the acute treatment of adults with major depression. *Journal of Consulting and Clinical Psychology*, 74(4), 658-670.

- Dozois, D.J. A., Dobson, K.S., & Ahnberg, J.L. (1998). A psychometric evaluation of the Beck Depression Inventory–II. *Psychological Assessment*, 10, 83-89.
- Elkin, I., Shea, M. T., Watkins, J. T., Imber, S. D. et al. (1989). National Institute of Mental Health Treatment of Depression Collaborative Research Program: General effectiveness of treatments. Archives of General Psychiatry, 46(11), 971-982.
- Elkin, I., Gibbons, R.D., Shea, M.T., Sotsky, S.M., Watkins, J.T., Pilkonis, P.A., &
 Hedeker, D. (1995). Initial severity and differential treatment outcome in the
 National Institute of Mental Health Treatment of depression collaborative
 research program. *Journal of Consulting and Clinical Psychology*, 63(5), 841.
- Fabrigar, L. R., Wegener, D. T., MacCallum, R. C., & Strahan, E. J. (1999). Evaluating the use of exploratory factor analysis in psychological research. *Psychological Methods*, 4(3), 272-299.
- First, M. B., Spitzer, R. L., Miriam, G., & Williams, J. B. W. (2002). Structured Clinical Interview for DSM-IV-TR Axis I Disorders, Research Version, Patient Edition. (SCID-I/P). New York: Biometrics Research, New York State Psychiatric Institute.
- Hamilton, M. (1960). A rating scale for depression. *Journal of Neurological Neurosurgical Psychiatry*, 23, 56-62.

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

- Hollon, S.D., Stewart, M.O., & Strunk, D. R. (2006). Enduring effects for cognitive behavior therapy in the treatment of depression and anxiety. *Annual Review of Psychology*, 57, 285-315.
- IAPT Programme (2007). The competencies required to deliver effective Cognitive and Behavioural Therapy for people with depression and with anxiety disorders.London: Department of Health.
- Jarrett, R.B., Schaffer, M., McIntire, D., Witt-Browder, A., Kraft, D., & Risser, R.C. (1999). Treatment of atypical depression with Cognitive Therapy or Phenelzine. *Archives of General Psychiatry*, 56, 431-437.
- Jarrett, R.B., Vittengl, J. R., Clark, L. A., & Thase, M. E. (2011). Skills of Cognitive Therapy (SoCT): A new measure of patients' comprehension and use. *Psychological Assessment, 23*, 578-586.
- Keller, M. B., Lavori, P. W., Friedman, B., Nielsen, E., et al. (1987). The longitudinal interval follow-up evaluation: A comprehensive method for assessing outcome in prospective longitudinal studies. *Archives of General Psychiatry*, 44(6), 540-548.
- Oliver, J. M., & Baumgart, E. P. (1985). The Dysfunctional Attitude Scale: Psychometric properties and relation to depression in an unselected adult population. *Cognitive Therapy and Research*, 9, 161-167.
- Overholser. (1991). The Socratic method as a technique in psychotherapy supervision. *Professional Psychology Research and Practice*, 22(1), 68.
- Overholser, J. C. (1993a). Elements of the Socratic method: I. Systematic questioning. *Psychotherapy: Theory, Research, Practice, Training, 30*(1), 67-74.

- Overholser, J. C. (1993b). Elements of the Socratic method: II. Inductive reasoning. *Psychotherapy: Theory, Research, Practice, Training, 30*(1), 75-85.
- Overholser, J. C. (1994). Elements of the Socratic method: III. Universal definitions. *Psychotherapy: Theory, Research, Practice, Training*, *31*(2), 286-293.
- Overholser, J. C. (1995). Elements of the Socratic method: IV. Disavowal of knowledge. *Psychotherapy: Theory, Research, Practice, Training, 32*(2), 283-292.
- Overholser, J. C. (2011). Collaborative empiricism, guided discovery, and the Socratic method: Core processes for effective cognitive therapy. *Clinical Psychology: Science and Practice, 18*(1), 62-66.
- Riskind, J.H., Beck, A.T., Brown, G., &Steer, R.A. (1987). Taking the measure of anxiety and depression: Validity of the reconstructed Hamilton scales. *Journal of Nervous and Mental Disease*, 175, 474-479.
- Sprinkle, S.D., Lurie, D., Insko, S.L., Atkinson, G., Jones, G.L., Logan, A.R., & Bissada, N.N. (2002). Criterion validity, severity cut scores, and test-retest reliability of the Beck Depression Inventory-II in a university counseling center sample. *Journal of Counseling Psychology*, 49, 381-385.
- Steer, R.A., Kumar, G., Ranieri, W.F., & Beck, A.T. (1998). Use of the Beck Depression Inventory-II with adolescent psychiatric *outpatients*. *Journal of Pychopathology* and Behavioral Assessment, 20, 127-137.
- Strunk, D. R., Brotman, M. A., & DeRubeis, R. J. (2006). Patients' responses to cognitive therapy: Examining therapist adherence and patient competencies.
 Presented in (R. J. DeRubeis) Cognitive therapy, medication, and the treatment of

depression: Outcomes, process, moderators, and mediators. Panel conducted at the Society for Psychotherapy Research Conference, Edinburgh, Scotland.

- Strunk, D. R., & DeRubeis, R. J. (2001). Cognitive therapy for depression: A review of its efficacy. *Journal of Cognitive Psychotherapy*, 15, 289-297.
- Strunk, D. R., DeRubeis, R. J., Chiu, A. W., & Alvarez, J. (2007). Patients' competence in and performance of cognitive therapy skills: Relation to the reduction of relapse risk following treatment for depression. *Journal of Consulting and Clinical Psychology*, 75, 523-530.
- Weissman, A. N. (1979). The Dysfunctional Attitude Scale: A validation study. Dissertation Abstracts International, 40, 1389B-1390B.
- Williams, J. B. (1988). A structured interview guide for the Hamilton Depression Rating Scale. Archives of General Psychiatry, 45, 742-747.

Appendix A: Patient Competencies Scale- Client Version

Patient Competencies Scale (PCS) – Client Version

Instructions.

The following questions ask you about how much (if at all) you have used some specific, strategies to cope with negative moods in <u>the last two weeks</u>.

There are several questions about "automatic thoughts" and ways you may have reacted to these thoughts. To understand these questions, it is important to know what is meant by the term automatic thoughts. Automatic thoughts are thoughts that come to mind without much conscious effort throughout your day-to-day life. At times, such thoughts may occur without your taking any special notice of them. This questionnaire deals primarily with negative automatic thoughts. For example, someone who got passed over for a promotion at work might think "I must have made too many mistakes. I was a bad employee."

With the understanding of automatic thoughts described above, please use the following scale to indicate how well it describes your thoughts, beliefs and behaviors over the last two weeks.

1	2	3	4	5	6	7
Not at all		Somewhat		Significantly		Completely

Please respond to every question honestly and to the best of your ability.

1 2 3 4 5 6 7	1. Rather than letting a challenge overwhelm me, I imagined how to break the challenge down, developed a plan, and worked on it step-by-step.
1 2 3 4 5 6 7	2. At times when my mood was at its lowest, I stepped back and recognized that my self evaluations were probably overly negative.
1 2 3 4 5 6 7	3. I was aware of some specific patterns in negative thinking that have tended to affect the way I interpret new situations.
1 2 3 4 5 6 7	4. I made an effort to evaluate my negative thoughts by considering just the facts.
1 2 3 4 5 6 7	5. I examined evidence from my past or present to more carefully consider whether my negative thoughts are accurate or not.
1 2 3 4 5 6 7	6. When my negative thoughts and emotions really bothered me, I had a specific action plan of things I could do to cope.

1 Not at all	2	3 Somewhat	4	5 Significantly	6	7 Completely
1 2 3 4 5 6 7	7. W and t	hen I had a neg ook time to eva	ative en lluate m	notional reaction, y negative though	I notice	d my negative thinking,
1 2 3 4 5 6 7	8. W else evalu	hen I evaluated would view the ate the situatio	a situat situatio n mysel	ion as negative, I n, and I used that f.	tried to to help	think of how someone me decide how to re-
1 2 3 4 5 6 7	9. I c	uestioned my on ative conclusion	original ons.	negative thoughts	s and ma	de an effort to develop
1 2 3 4 5 6 7	10. I abou	recognized tha t myself.	t negativ	ve feelings are rel	lated to r	negative thoughts I have
1 2 3 4 5 6 7	11. V think	When I found m ing and worked	iyself up l to deve	oset about someth elop a more balan	ing, I too leed viev	ok note of what I was v.
1 2 3 4 5 6 7	12. I relati	recognized that	t beliefs o longer	I formed on the be applicable in	basis of j the same	past events and way today.
1 2 3 4 5 6 7	13. I help	have been reco to maintain my	ognizing depress	that inaccurate, r	negative	thoughts and judgments
1 2 3 4 5 6 7	14. I that t	have been awa end to fuel my	re of spo negative	ecific patterns in e emotions.	my nega	tive thinking – beliefs
1 2 3 4 5 6 7	15. I abou	took note of ar t myself (e.g., i	nd wrote n a jour	down specific ne nal or diary).	egative tl	houghts or judgments
1 2 3 4 5 6 7	16. V remin happ	When I found m nded myself tha en.	nyself wat the co	orrying that some nsequences migh	ething ba t not be s	d would happen, I so terrible even if it did
1 2 3 4 5 6 7	17. V that i	When I made an my mood would	effort t get bet	o correct my negater.	ative thin	nking, I was confident
1 2 3 4 5 6 7	18. V woul	When I was ups d be likely to in	et, I mae nprove	de an effort to eng my mood.	gage in e	njoyable activities that
1 2 3 4 5 6 7	19. V gave	When I have fel me a sense of a	t down, accompl	I engaged in activitishment (to try to	vities that help my	tt were enjoyable or 7 mood).
1 2 3 4 5 6 7	20. I moo	have been cont d would improv	fident th 'e.	at if I made an ef	fort to b	e less pessimistic, my

1 Not at all	2 3 Somewhat	4	5 Significantly	6	7 Completely
1 2 3 4 5 6 7	21. I caught myself evaluated the situation	thinking on.	negatively, reco	gnized th	ne negative bias, and re-
1 2 3 4 5 6 7	22. I often caught m to develop more rati	yself thin onal viev	nking in an irrati vs.	ional way	and I actively worked
1 2 3 4 5 6 7	23. When I had a tas effort to break the ta	sk that I i sk up int	might have had t o smaller parts.	rouble u	ndertaking, I made an
1 2 3 4 5 6 7	24. I took time to reparts of my day.	view spe	cific thoughts I	had durin	g the most upsetting
1 2 3 4 5 6 7	25. When I blamed a consider other factor	myself for s that ma	or something bac ay have been inv	l that hap volved.	opened, I took time to
1 2 3 4 5 6 7	26. I believed that w would probably help	orking to ease my	o more carefully depression.	evaluate	my negative thoughts
1 2 3 4 5 6 7	27. I noticed specific	c automa	tic thoughts as t	hey occu	rred.
1 2 3 4 5 6 7	28. Rather than avoid developed a solution	ding a di 1, and fol	ifficult decision, lowed through v	I weighe vith the d	ed my options, lecision I made.
1 2 3 4 5 6 7	29. When I got upse that my negative tho	et, I took aughts mi	time to step back ght be inaccurat	k from a æ.	situation and consider
1 2 3 4 5 6 7	30. When I had a too to be assertive in add encounter, and initia	ugh inter dressing ted a cor	personal issue to the issue, tried to aversation on the	o address o anticipa e topic.	, I thought through how ate obstacles I might

Appendix B: Patient Competencies Scale- Therapist Version

Patient Competencies Scale (PCS) – Therapist version

Instructions.

The following questions ask you about your client and his or her ability, independent use and understanding of, and frequency of use for major components of Cognitive Therapy.

Please rate your client across three domains (Behavioral Activation, Automatic Thoughts, and Core Beliefs) on three characteristics: their ability, independence, and frequency of using these strategies.

Behavioral Activation (BA)

Ability: Use the following scale to rate the participant's ability in the BA domain:

0	1	2	3	4	5	6
Not at all		Some Ability	Coi	nsiderable Ab	ility	Extensive Ability

<u>Independence</u>: Use the following scale to rate the participant's independence in using BA strategies:

0	1	2	3	4	5	6
No		Some		Considerable		Extensive
Independence		Independence		Independence		Independence

<u>Frequency</u>: Use the following scales to rate the frequency with which the participant has been using BA strategies:

0	1	2	3	4	5	6
Not at all		Somewhat		Considerably		Extensively

Automatic Thoughts (AT)

Ability: Use the following scale to rate the participant's ability on the AT domain:

0	1	2	3	4	5	6
Not at all		Some Ability	Co	nsiderable At	oility	Extensive Ability

<u>Independence</u>: Use the following scale to rate the participant's independence in using AT strategies:

0	1	2	3	4	5	6
No		Some		Considerable		Extensive
Independence		Independence		Independence		Independence

<u>Frequency</u>: Use the following scales to rate the frequency with which the participant has been using AT strategies:

0123456Not at allSomewhatConsiderablyExtensively

Core Beliefs (CB)

Ability: Use the following scale to rate the participant's ability on the CB domain:

0	1	2	3	4	5	6
Not at all		Some Ability	Co	nsiderable Abi	lity	Extensive Ability

<u>Independence</u>: Use the following scale to rate the participant's independence in using CB strategies:

0	1	2	3	4	5	6
No		Some		Considerable		Extensive
Independence		Independence		Independence		Independence

<u>Frequency</u>: Use the following scales to rate the frequency with which the participant has been using CB strategies:

0	1	2	3	4	5	6	
Not at all		Somewhat		Considerably		Extensively	

Appendix C: Tables

	Inta	ke			
	PCS-C		$\begin{array}{c} \text{PCS-C} \\ (n = 43) \end{array}$	PCS	5-T
	r	n	r	r	n
WOR- Quality	.30*	63	.37*	.35*	43
WOR- Total	.36**	63	.30*	.29*	43
DAS	17	64	30*	27†	40
BDI-II	39**	63	49**	59**	40
HRSD	22†	65	46**	.54**	40

Table 1: Correlations of pre and post treatment patient competencies with measures of CT skills and depression symptom severity

Note: The PCS-T was administered only at the post treatment assessment. † p < .10; * p < .05; ** p < .01 Table 2: Correlations of PCS residualized change with pre to post treatment change in alternative compensatory skills measures and slopes of change of depressive severity

Correlations with Residualized Change in DAS and WOR
--

PCS-C Residualized Change		PCS-T Residualized Chang		ge
r	n	r	п	
.34*	40	.34*	39	
.27†	40	.27†	40	
31*	42	32*	39	
	PCS-C Residua .34* .27† 31*	r n .34* 40 .27† 40 31* 42	PCS-C Residualized Change PCS-T Residu r n r .34* 40 .34* .27† 40 .27† 31* 42 32*	PCS-C Residualized Change PCS-T Residualized Change r n r n .34* 40 .34* 39 .27† 40 .27† 40 31* 42 32* 39

HLM analysis of Change in BDI and HRSD

PCS-C Residualized Change PCS-T Residualized Change

	r	t	n	r	t	п
BDI change	50**	-3.06	42	32*	-2.17	41
HRSD change	48**	-3.58	42	26†	-1.75	41

Note. PCS-T residualized change scores calculated using PCS-C intake scores as PCS-T data were not recorded at intake. $\dagger p < .10$; * p < .05; ** p < .01

	BDI-II	HRSD			
	Mode	1 1			
Intercept	27.90 (1.53)**	19.96 (.69)**			
Time	-4.50 (.48)**	71 (.06)**			
WOR Quality	1.96 (1.99)	.71 (.90)			
PCS-C	.08 (.06)	.03 (.03)			
PCS-T	15 (.15)	06 (.07)			
WOR Quality x Time	17 (.64)	11 (.08)			
PCS-C x Time	04 (.02)†	01 (.00)*			
PCS-T x Time	03 (.05)	00 (.01)			
	Model 2				
Intercept	27.62 (1.53)**	19.73 (.67)**			
Time	-4.42 (.50)**	69 (.06)**			
WOR Quality	2.21 (1.97)	.67 (.88)			
PCS-T	07 (.13)	03 (.06)			
WOR Quality x Time	33 (.65)	14 (.08)			
PCS-T x Time	08 (.04)†	01 (.01)			
	Model 3				
Intercept	27.37 (1.47)**	19.85 (.74)**			
Time	-4.40 (.47)**	69 (.06)**			
WOR Quality	1.72 (1.94)	.03 (.03)			
PCS-C	.04 (.05)	.76 (.97)			
WOR Quality x Time	14 (.63)	11 (.08)			
PCS-C x Time	04 (.02)*	01 (.00)*			

Table 3: Parameter estimates (and Standard Errors) for Mixed Models examining residualized change in WOR Quality and PCS as predictors of HRSD and BDI-II change

Note. † p < .10; * p < .05; ** p < .01

	BDI-II	HRSD
	Model 1	
Intercept	27.92 (1.51)**	19.98 (.67)**
Time	-4.50 (.47)**	71 (.06)
WOR Total	.18 (.13)	.09 (.06)
PCS-C	.08 (.06)	.03 (.03)
PCS-T	15 (.15)	07 (.07)
WOR Total x Time	04 (.04)	01 (.01)*
PCS-C x Time	04 (.02)†	01 (.00)*
PCS-T x Time	03 (.05)	.00 (.01)
	Model 2	
Intercept	27.63 (1.51)**	19.74 (.66)**
Time	-4.43 (.49)**	69 (.06)**
WOR Total	.19 (.13)	.09 (.06)
PCS-T	07 (.12)	01 (.01)
WOR Total x Time	05 (.04)	01 (.01)*
PCS-T x Time	07 (.04)†	01 (.01)
	Model 3	
Intercept	27.35 (1.46)**	19.85 (.72)**
Time	-4.40 (.46)**	69 (.06)**
WOR Total	.15 (.13)	.08 (.06)
PCS-C	.04 (.05)	.03 (.02)
WOR Total x Time	04 (.04)	01 (.01)†
PCS-C x Time	04 (.02)*	01 (.00)**

Table 4: Parameter estimates (and Standard Errors) for Mixed Models examining residualized change in WOR Total and PCS as predictors of HRSD and BDI-II change

Note. $\dagger p < .10$; * p < .05; ** p < .01