CHANGE IN PSYCHOSOCIAL FUNCTIONING DURING

COGNITIVE THERAPY FOR DEPRESSION

APPROVED BY SUPERVISORY COMMITTEE

Robin B. Jarrett, Ph.D. (Chair)

Thomas Carmody, Ph.D.

Bert Moore, Ph.D.

Celette Skinner, Ph.D.

Deborah J. Wiebe, Ph.D., M.P.H.

DEDICATION

If this dissertation has taught me anything, it is that success cannot happen alone. Over the past two and a half years, many trusted mentors, friends, and family members have helped me complete this project. I want to first thank my dissertation supervisor, Dr. Jarrett for her continual support and encouragement. She not only helped me mature as a researcher, but she also took time to show me how to be a better colleague, teacher, and clinician. I don't know how she did it, but she always made me feel like my dissertation was her most important project, even though I knew she had a thousand other demands.

I also want to thank the members of my dissertation committee for their tireless efforts to improve this work. Specifically, I want to thank Dr. Carmody for his statistical assistance that saw me through the confusing labyrinth of structural equation modeling; Dr. Moore for his focus on the "grander scheme of things", calming presence, and kind words; Dr. Skinner for her enthusiasm, gentle reminders to get "that SCT paper"

published, and for making the world of academia look fun; Dr. Weibe for her willingness to join the committee and see the project through; and Dr. Basco for her helpful feedback.

I want to thank my classmates for being there through thick and thin and reminding me to not take myself so seriously. All those hoops we jumped through to graduate have provided many fond memories that will make me laugh in the years to come…just not right now☺

I could not have completed this dissertation without the support of my loving family. Growing up on a small, country farm, my mother and father instilled in me a work ethic and love of learning that have brought me to this point. I thank my sister and brother for making me laugh and "keeping it real" even though I have at times dropped off the face of the earth. Also, my mother and father-in-law have been stabilizing influences in my life that have helped me stay afloat. I am eternally grateful to each of you.

Lastly, I want to dedicate this work to my beautiful wife and adorable daughter and son. They more than any other have known the sacrifice involved in this work and, despite the countless late nights at the office, have been a source of unconditional love. Little did they know that it was thoughts of them that got me through each day, thoughts of going home and seeing their smiling faces, jumping on the bed, or snuggling up for movie night. They have been my foundation, without which I would have fallen years ago. I owe this achievement to them.

CHANGE IN PSYCHOSOCIAL FUNCTIONING DURING COGNITIVE THERAPY FOR DEPRESSION

by

TODD WILSON DUNN

DISSERTATION

Presented to the Faculty of the Graduate School of Biomedical Sciences

The University of Texas Southwestern Medical Center at Dallas

In Partial Fulfillment of the Requirements

For the Degree of

DOCTOR OF PHILOSOPHY

The University of Texas Southwestern Medical Center at Dallas

Dallas, Texas

December, 2008

CHANGE IN PSYCHOSOCIAL FUNCTIONING DURING

COGNITIVE THERAPY FOR DEPRESSION

TODD WILSON DUNN, M.S.

The University of Texas Southwestern Medical Center at Dallas, 2008

ROBIN BETH JARRETT, Ph.D.

Major Depressive Disorder (MDD) is a highly prevalent and recurrent disorder that impairs peoples' work, relationships, and leisure activities. Cognitive Therapy (CT) improves this impairment in psychosocial functioning in adults with MDD, but questions remain as to how improvements occur both independently and in relation to depressive symptoms. To address this issue, the current study developed a theoretical framework based on social cognitive theory to conceptualize change in psychosocial functioning during CT and tested it with structural equation modeling. Using data from 470 patients undergoing acute-phase CT (A-CT) for MDD, results showed that: a) change in psychosocial functioning and depressive symptom severity occurred independently of

each other, b) change in psychosocial functioning during the first month of A-CT partially mediated change in depressive symptom severity from treatment baseline to week seven of A-CT, and c) psychosocial functioning at week seven of A-CT significantly predicted subsequent depressive symptom severity. In terms of the theoretical framework, results suggested that when people with MDD were exposed to an environmental stimuli (i.e., acute-phase CT), change in their behavior (i.e., psychosocial functioning) partially mediated change in personal factors (i.e., depressive symptom severity) and not vice versa. By disentangling the sequence of change in psychosocial functioning and depressive symptom severity, this study pushed the field one step closer to understanding how A-CT treats the impairment in psychosocial functioning associated with MDD.

TABLE OF CONTENTS

PRIOR PUBLICATIONS

Dunn, T. W. & Jarrett, R. B. (in press). Psychosocial functioning in depression. In R. Ingram (Ed.), *The International Encyclopedia of Depression*.

Vittengl, J. R., Clark, L. A., Dunn, T. W., & Jarrett, R. B. (2007). Reducing relapse and recurrence in unipolar depression: A comparative meta-analysis of cognitive therapy's effects. *Journal of Consulting and Clinical Psychology, 75,* 475-488.

Dunn, T. W., Smith, T. B., & Montoya, J. (2006). Multicultural competency instrumentation: A review and analysis of reliability generalization. *Journal of Counseling and Development*, *84,* 471-482.

Smith, T. B., Constantine, M. G., Dunn, T. W., Dinehart, J. M., & Montoya, J. A. (2006). Multicultural education in the mental health professions: A meta-analytic review. *Journal of Counseling Psychology, 53,* 132-145.

Dunn, T. W., Burlingame, G. M., Walbridge, M., Smith, J., & Crum, M. J. (2005). Outcome assessment for children and adolescents: Psychometric validation of the Youth Outcome Questionnaire 30.1 (YOQ®-30.1). *Clinical Psychology & Psychotherapy, 12,* 388-401.

Burlingame, G. M., Dunn, T. W., Chen, S., Lehman, A., Axman, R., Earnshaw, D., & Rees, F. M. (2005). Selection of outcome assessment instruments for inpatients with severe and persistent mental illness**.** *Psychiatric Services, 56,* 444-451.

Earnshaw, D., Rees, F. M., Dunn, T. W., Burlingame, G. M., & Chen, S. (2005). Implementing a multisource outcome assessment protocol in a state psychiatric hospital: A case study. *Psychiatric Services, 56,* 411-413.

Burlingame, G. M., Dunn, T. W., Hill, M., Cox, J., Wells, G., Lambert, M., & Brown, G. S. (2004). *Administration and Scoring Manual for the Y-LSQ^{TM} 2.0 (Youth Life Status Questionnaire-Version 2.0).* Stevenson, MD: American Professional Credentialing Services.

Robinson, P. W., Robinson, M., & Dunn, T. W. (2003). STEP parenting: A review of the research. *Canadian Journal of Counseling, 37,* 270-278.

LIST OF FIGURES

LIST OF TABLES

LIST OF APPENDICES

CHAPTER ONE
Introduction

During their lifetime, 16% of Americans will be diagnosed with Major Depressive Disorder (MDD), a costly, chronic, and disabling disorder (Kessler, Berglund, Demler, Jin, & Walters, 2005). To be diagnosed with MDD, a person must suffer from both depressive symptoms, including negative cognitions, emotions, and somatic events, as well as impairment in psychosocial functioning (American Psychiatric Association, 2000). Although the impairment in psychosocial functioning associated with MDD rivals that of chronic diseases (Hays, Wells, Sherbourne, Rogers, & Spritzer, 1995), few have investigated when and how it changes during treatment. Instead, researchers investigating the treatment of MDD focus almost exclusively on changes in depressive symptoms.

Defined as an individual's ability to function in his or her environment (Paykel, 1999), impairment in psychosocial functioning is linked with the onset and persistence of depressive symptoms (e.g., Moos & Cronkite, 1999; Prince, Harwood, Thomas, & Mann, 1998), poor response to treatment (e.g., Hirschfeld et al., 1998; Hoencamp, Haffmans, & Duivenvoorden, 1998; Sotsky et al., 1991), and higher incidences of relapse and recurrence (e.g., Leon et al., 1999; Rodriguez, Bruce, Pagano, & Keller, 2005; Vittengl, Clark, & Jarrett, 2008). Given the above, some claim that effective acute-phase interventions for MDD should not only reduce depressive symptom severity, but also return individuals to premorbid levels of functioning (e.g., Bakish, 2001; Keller, 2003; Thase, 2003; Trivedi, 2001).

When provided during the acute phase, or the first 12 weeks of treatment (Rush et al., 2006), cognitive therapy (CT; Beck, Rush, Shaw & Emery, 1979) effectively reduces depressive symptom severity (Butler, Chapman, Forman, & Beck, 2006) and impairment in psychosocial functioning (e.g., Imber et al., 1990; Vittengl, Clark, & Jarrett, 2004). Despite these encouraging results, however, it remains unclear when and how these two factors change (e.g., Garratt, Ingram, Rand, & Sawalani, 2007; Hirschfeld et al., 2002). Due to the centrality of psychosocial functioning in the treatment and diagnosis of MDD, new knowledge will be created if researchers can address when and how psychosocial functioning changes during acute-phase CT (A-CT), both independently and in relation to depressive symptom severity.

When confronting these questions, two issues arise. First, the constructs of psychosocial functioning and depressive symptom severity correlate highly (e.g., Judd, Akiskal, et al., 2000). In fact, some question the utility of measuring psychosocial functioning to diagnose or track treatment outcomes with depression because it may add little information beyond that already provided by measures of symptom severity (Rush et al., 2006; Spitzer & Wakefield, 1999). Consequently, before investigating when and how psychosocial functioning changes during treatment for MDD, researchers need to provide evidence to suggest depressive symptom severity and psychosocial functioning are unique constructs. Researchers can do this by controlling for sources of shared variability in their analyses, thereby isolating variance unique to each construct.

Second, the field lacks a clear understanding of how to conceptualize the changes that occur in psychosocial functioning during A-CT. To meet this need, a theoretical framework should be developed, or an existing theory applied, to explain how specific

therapeutic factors bring about change in psychosocial functioning in individuals suffering from MDD. This framework should describe how psychosocial functioning differs from, overlaps with, or is influenced by other elements of psychopathology (e.g., environment, biology, cognition, etc.), further differentiating its role in the course and treatment of MDD. Also, if this framework is to allow researchers and clinicians to better account for changes in psychosocial functioning during depression-specific treatment as a whole, it needs to be heuristically useful in research where distinct types of treatment are provided (e.g., other forms of psychotherapy or pharmacotherapy).

The proposed dissertation addresses the above issues in an attempt to better understand when and how changes occur in psychosocial functioning, using data from an open trial of 470 individuals receiving A-CT for recurrent MDD. The hypotheses and interpretations of this dissertation rest upon a theoretical framework based on social cognitive theory (Bandura, 1986, 1997). Social cognitive theory is chosen for the basis of this framework because it is: a) comprehensive in terms of content covered by depression-specific theories and therapeutic techniques, b) parsimonious in that it avoids unnecessary complexity when incorporating important constructs in the field, c) flexible in that it accounts for inter-correlation between constructs along reciprocal casual pathways, and d) predictive, as it is based on a body of literature supporting hypothesis testing.

Since all participants in this dissertation received A-CT, cognitive theory (Beck et al., 1979) was also considered as a potential theoretical framework. Social cognitive theory and cognitive theory are similar in that they both claim cognition mediates the impact of the environment on behavior (Bandura, 1997; Clark, Beck, & Alford, 1999).

The main differences are that cognitive theory focuses on cognition as the primary mechanism of change in emotion and behavior, and it is largely phenomenological, consisting of a person's perceptions of his or her environment, future, and self (Beck & Weishaar, 1993). On the other hand, social cognitive theory claims that environmental (e.g., life events, socioeconomic status, education, familial structures, etc.), personal (i.e., cognitive, affective, and biological events), and behavioral factors all interact with each other along reciprocal causal pathways (Bandura, 1986, 1997, 1999, 2001). So, depending on the circumstances, changes in each of these factors can hypothetically cause or mediate changes in the others. In addition, social cognitive theory places an equal emphasis on subjective perceptions and objective observations when identifying each factor.

In order to apply social cognitive theory to better understand how change occurs in psychosocial functioning during A-CT, this study operationalizes the situation or environment to be A-CT and holds it constant for all participants. In addition, personal and behavioral factors are operationalized to be depressive symptoms (i.e., negative emotions, cognitions, and somatic events) and psychosocial functioning, respectively. So, in terms of social cognitive theory, this dissertation explores how changes occur in behavior (i.e., psychosocial functioning) within the context of A-CT, both independently and in conjunction with personal factors (i.e., depressive symptoms). While social cognitive theory allows for some correlation between constructs through reciprocal causal pathways, prior research shows changes in depressive symptom severity largely account for changes in psychosocial functioning during depression-specific treatment (e.g., Hirschfeld et al., 2002; Vittengl et al., 2004). As a result, it is predicted that changes in

depressive symptom severity will partially mediate changes in psychosocial functioning during A-CT for MDD and not vice versa. To test this hypothesis, the proposed dissertation utilizes structural equation modeling. The benefit of structural equation modeling is that it can control for shared variability between measures of psychosocial functioning and depressive symptom severity to better determine what change occurs in variance unique to each construct, both independently and in relation to each other.

The results from this study may have implications on how the field, not just those practicing CT, understands the treatment of impaired psychosocial functioning in individuals with MDD. For example, by clarifying when and how change occurs in psychosocial functioning in relation to change in depressive symptom severity, this study works to disentangle the sequence of change in symptoms and functioning during A-CT. Specifically, study results may suggest that when people are exposed to a particular environment (i.e., A-CT) change in their behavior (i.e., psychosocial functioning) is mediated by changes in personal factors (i.e., depressive symptoms). With this knowledge, researchers and clinicians come one step closer to understanding how A-CT treats MDD. Also, if the theoretical framework employed in this dissertation to conceptualize change in psychosocial functioning is substantiated, researchers can examine the extent to which it generalizes to other psychosocial and pharmacological treatments for MDD.

In addition to furthering the field's conceptual understanding, this study may also have practical implications for the measurement and treatment of recurrent MDD. For example, if researchers know what is unique about the construct of psychosocial functioning, especially compared to the construct of depressive symptom severity, they

may be better able to design instruments to assess it. Such technology would improve efforts to assess psychosocial functioning during diagnostic and treatment outcome evaluations. In addition, with the information this study provides, clinicians may be better able to inform their clients about when and how they are likely to experience improvements in their psychosocial functioning. Therefore, in the context of adequate communication and informed consent between practitioner and patient, these results may help reduce the risk of premature termination, as both parties will have a better understanding of how long it takes and under what circumstances change occurs in psychosocial functioning.

Specific Aims

To set the stage for the primary aim, preliminary analyses examine the degree to which measures of psychosocial functioning and depressive symptom severity correlate and change during A-CT. It was expected that these preliminary analyses would corroborate past research and show the constructs of psychosocial functioning and depressive symptom severity correlate and improve significantly during A-CT. The question targeted by this dissertation's primary aim asks what happens to that portion of variance unique to measures of psychosocial functioning during A-CT for MDD. Specifically, after controlling for the variance shared between measures of psychosocial functioning and depressive symptom severity, **the Primary Aim of this dissertation goes beyond existing literature to:**

A1) determine the extent to which change in variance unique to measures of depressive symptom severity mediates change in variance unique to measures of psychosocial functioning.

To accomplish this, a structural equation model was fit with data to account for: a) covariation between measures of psychosocial functioning and depressive symptom severity, b) shared error variance across repeated measurements, c) change in depressive symptom severity that occurs independently of change in psychosocial functioning, d) change in psychosocial functioning that occurs independently of change in depressive symptom severity, and e) potential mediating relationships between constructs. In this study, a mediating variable: a) occurs between measurements of the independent and dependent variables, b) correlates with the independent variable, and c) accounts for part or all of the relationship between the independent and dependent variables (Wilson, Fairburn, Agras, Walsh, & Kraemer, 2002).

Based on previous research (e.g., Hirschfeld et al., 2002; Lenderking et al., 1999; Vittengl et al., 2004), it is predicted that changes in variance unique to measures of depressive symptom severity partially mediate changes in variance unique to measures of psychosocial functioning and not vice versa. It is also predicted that the variance unique to each construct changes independently over the course of A-CT. Potential implications of this finding are discussed in terms of the theoretical framework developed from social cognitive theory to explain changes in psychosocial functioning and the therapeutic process in general.

The preliminary aims of the dissertation seek to:

B1) psychometrically validate the indices of psychosocial functioning and depressive symptom severity, which are used to operationalize each construct in the primary aim, and

B2) determine the extent to which psychosocial functioning and depressive symptom severity improve after exposure to A-CT compared to pre-treatment baseline.

Results from these preliminary aims facilitate interpretation of the primary aim, because one first needs to show that change occurs and is measured reliably and validly before trying to explain how it occurs.

CHAPTER TWO
Qualitative Review of the Literature

Major Depression Affects Public Health Adversely

Major Depressive Disorder (MDD) is a highly prevalent disorder that costs the United States 83.1 billion dollars (Greenberg et al., 2003) and is implicated in over 15,500 suicide fatalities every year (Henriksson et al, 1993; Minino, Heron, & Smith, 2006; Oquendo, et al 2001). Researchers estimate that one in six Americans suffer from MDD at some point in their life (Kessler et al., 2005), including one in seven individuals in primary medical settings (Ansseau et al., 2004). Given that physicians fail to recognize depression half the time (Bhugra, & Mastrogianni, 2004), these estimates could easily be doubled, meaning approximately one in every four patients in primary medical settings has comorbid depression.

The adverse public health impact of MDD is attributable in part to its recurrent, and potentially fatal nature. For instance, Mueller et al (1999) found that 85% of those who recover from a Major Depressive Episode (MDE) during acute-phase treatment can expect to experience a recurrent episode of depression. Furthermore, of those who experience a recurrent MDE, the likelihood that they will experience another MDE increases exponentially with each successive recurrence, with a 90% chance of recurrence after the third MDE (American Psychiatric Association, 2000; Solomon et al., 2000).

Interest in the Construct of Psychosocial Functioning Reappears

Interest in the construct of psychosocial functioning first peaked in the 1960's and 70's, when researchers began emphasizing the social context of mental illness and mental health care shifted from psychiatric hospitals to outpatient, community health centers (Weissman, 2000). During that time, researchers encouraged clinicians to focus on improving their patients' functioning at home, work, and in the community (Brayfield, 1965; McNair & Lorr, 1964). Concurrent research investigated the role that psychosocial functioning played in predicting treatment response (e.g., Rounsaville, Weissman, & Prusoff, 1981; Simons, Levine, Lustman, & Murphy, 1984) and relapse or recurrence in individuals with MDD (e.g., Paykel & Weissman, 1973; Tanner, Weissman, & Prusoff, 1975). In response to these efforts, governing bodies developed separate classification systems to assess psychosocial functioning (i.e., American Psychiatric Association, 1980; World Health Organization, 1980), thereby highlighting the importance of functional impairment in the diagnosis of mental illness. However, interest in social functioning "waned after these initial efforts" (p. 34; Weissman, 2000).

Over the past decade, researchers have again focused on the functional impairments associated with MDD for four primary reasons. First, many researchers have determined that the current definitions of treatment response are inadequate because they rely exclusively on assessments of symptomatic change (Bakish, 2001; Keller, 2003; Kennedy, Eisfeld, & Cooke, 2001; Thase, 2003; Trivedi, 2001). Zimmerman et al. (2006) emphasized this point, claiming changes in depressive symptoms and psychosocial functioning should both be taken into account in clinical trials because both

are required to receive a diagnosis of MDD. Second, given the increasing importance of consumer satisfaction in the managed health care environment (Healy & McMonagle, 1997), researchers are taking notice that clients consider their functional status an important area of treatment outcome (Hasler & Moergeli, & Schnyder, 2004; Mee & Sumsion, 2001; Zimmerman et al., 2006). Third, mental health care providers increasingly rely on measures of psychosocial functioning to justify new, often more expensive treatment strategies to third party payers (Weissman, Olfson, Gameroff, Feder, & Fuentes, 2001). For example, in the case of reboxetine, a noradrenaline reuptake inhibitor, researchers went to such lengths as to create a new measure of psychosocial functioning to show their medication was superior to existing anti-depressant medications (Bosc, Dubini, & Polin, 1997; Dubini, Bosc, & Polin, 1997). Lastly, Greenberg et al. (2003) showed that 62% of the economic burden of depression was due to indirect costs (e.g., lost work productivity, absenteeism, etc.), while 31% was due to the direct costs of treatment. Therefore, it appears that researchers can have the greatest impact on the "real world" burden of depression by investigating how the associated impairment in psychosocial functioning is best alleviated.

The Field Needs a Consensus Definition of Psychosocial Functioning

To be diagnosed with MDD, a person's depressive symptoms must cause them "clinically significant distress or impairment in social, occupational, or other important areas of functioning" (p. 356; American Psychiatric Association, 2000). Despite including functional impairment as a criterion for MDD, however, diagnostic manuals

offer little else to help clinicians define or quantify the construct. In addition, over 30 measures of psychosocial functioning exist (e.g., Bech, 2005; Bosc, 2000; Hilsenroth et al., 2000; Kennedy et al., 2001; Weissman, Sholomskas, & John, 1981; Wiersma, 1996), each with a unique conceptualization of the construct (Weissman et al., 2001). Further complicating matters, researchers have created different terms to refer to the construct of psychosocial functioning (e.g., quality of life, psychosocial functioning, social functioning, social adaptation, social adjustment, etc.), sometimes using them interchangeably (e.g., Kennedy et al., 2001; Papakostas, Petersen, Mahal, et al., 2004).

Unfortunately, the inconsistent use of terminology and existence of divergent diagnostic and assessment techniques hinder the field by limiting comparisons across studies. If the field is to ever agree on how impairments in psychosocial functioning are assessed and diagnosed, researchers and clinicians need a consensus definition of the construct and a standardized, psychometrically sound source of measurement. In this dissertation, the internal consistency reliability, construct validity, and sensitivity to change of five measures of psychosocial functioning are compared (see Appendix G). These analyses help the field determine how the construct of psychosocial functioning is best measured.

How is Psychosocial Functioning Currently Defined?

Parsons (1958) was the first to define psychosocial functioning as a person's performance in the roles or tasks for which he or she has been socialized. Almost half a century later, Hirschfeld et al. (2000) defined psychosocial functioning as a person's

performance in and satisfaction with "normal social roles", including occupational, interpersonal, and recreational domains. Though largely similar, more recent definitions of psychosocial functioning have emphasized the need for both objective and subjective perspectives when measuring the construct (e.g., Hirschfeld et al., 2000; Paykel, 1999). While objective estimates are usually made by third-party observers (e.g., clinician, parent, etc.) who quantify or describe the behaviors designed to fulfill social obligations, subjective estimates are self-report and involve quantifying the extent to which role fulfillment meets personally relevant standards, which can include both satisfaction with aspects of the role itself and the person's performance within it. In advocating the inclusion of the subjective perspective, Endicott, Nee, Harrison, and Blumenthal (1993) said the capacity to "feel satisfied with one's ability to function in multiple life activities has been recognized to be as important to many patients as is the absence of signs and symptoms used to define the disorders" (p. 321).

Table 1 provides a heuristic for defining and measuring psychosocial functioning (Dunn & Jarrett, in press; available in Appendix A). With occupational functioning, a person rates his or her performance at work, school, and/or home along levels of investment, workload, or productivity. A person's interpersonal functioning refers to his or her fulfillment of responsibilities as a partner, spouse, parent, family member, extended relative, friend, and/or neighbor. Within these roles, consideration is given to the frequency, quality, and content of relationships, including sexual functioning. Finally, a person's recreational functioning includes behavior that may involve "leisure" activities, or activities unrelated to societal "duties". To measure this particular domain,

people report on the quantity and quality of recreational activities and hobbies (e.g., sports, gardening, reading, watching TV, going to movies, church, etc.).

Table 1

Measuring Psychosocial Functioning using both Objective and Subjective Perspectives

Measurement of Psychosocial Functioning			
Objective assessment of performance in each role ↑	Interpersonal Functioning with:	Occupational Functioning at:	Recreational Functioning in:
	Spouse/Partner Parents Children Siblings Extended family Neighbors Friends	Work School Home	Hobbies Exercise Religious service Community service Other recreational activities
↓	← Satisfaction with content and performance in each role →		

Note. This table can also be found in Dunn & Jarrett (in press)

Despite differences across domains, social roles and the evaluation of functioning therein may overlap and/or conflict. For example, a father who plays with his children after work can be functioning in both the recreational and interpersonal domains. Likewise, an employee who aggressively (but rudely) promotes her own agenda items in a business conference may perceive herself as functioning well within the occupational domain, while her colleagues perceive her as having poor interpersonal skills. In sum, a single event can be used to rate functioning in more than one domain, and the quality of

the functioning that occurred during that event can vary depending on the domain and rater.

How Can Change in Psychosocial Functioning during A-CT for MDD be Conceptualized?

Even though definitions have existed since the 1950's, few efforts have been made to establish a theoretical framework for psychosocial functioning (Bosc et al., 1997; Weissman et al., 1981), by explaining how it relates to other important constructs in the field (e.g., environment, genetics, emotions, cognitions, etc.). This may be because attempts to define psychosocial functioning have been driven by the need to develop instruments to assess community adjustment and provide a social context for clinical syndromes (Weissman, 2000). As a result, researchers often define the construct as a list of behaviors or attributes assumed to capture a person's ability to function in society (e.g., Gurland, Yorkston, Stone, Frank, & Fleiss, 1972; Weissman & Bothwell, 1976). Therefore, it can be said that the field's understanding of psychosocial functioning is based on experiential or clinical data rather than a clear theoretical framework.

In this section, a theoretical framework is developed based on social cognitive theory (Bandura, 1986, 1997) to explain how psychosocial functioning relates to other important constructs in the field, including the environment, biology, emotion, and cognition. Specifically, this section: a) explains why social cognitive theory is chosen as the basis of the theoretical framework, b) describes the basic tenants of social cognitive theory, c) conceptualizes psychosocial functioning within the context of social cognitive

theory, d) predicts how this theoretical framework will account for changes in psychosocial functioning during A-CT, and e) discusses implications of study findings for social cognitive theory and the treatment of MDD with A-CT.

While this dissertation examines psychosocial functioning in the context of depression, the construct is defined and conceptualized in a general sense. The purpose for so doing is to encourage future research on how psychosocial functioning pertains to the development, course, and treatment of diverse psychopathology and disease processes. At the same time, researchers can still seek to replicate this study's hypotheses regarding psychosocial functioning and depression with other psychotherapies (e.g., interpersonal psychotherapy) or treatments modalities (e.g., pharmacotherapy).

Why Conceptualize Psychosocial Functioning Using Social Cognitive Theory?

In this dissertation, social cognitive theory (Bandura, 1986, 1997) is used to conceptualize psychosocial functioning and better understand how it changes during A-CT for MDD. This particular theory is chosen for four reasons: comprehensiveness, parsimony, flexibility, and predictability.

First, social cognitive theory incorporates environmental, biological, cognitive, emotional, and behavioral factors in a comprehensive model of human experience (Bandura, 2001). Many theories offer frameworks for understanding changes in psychosocial functioning during depression-specific treatment, but most are focused on a specific area of interest. For instance, cognitive, psychoanalytic, psychopharmacological, and behavioral theories of depression target maladaptive cognitions, internalized feelings

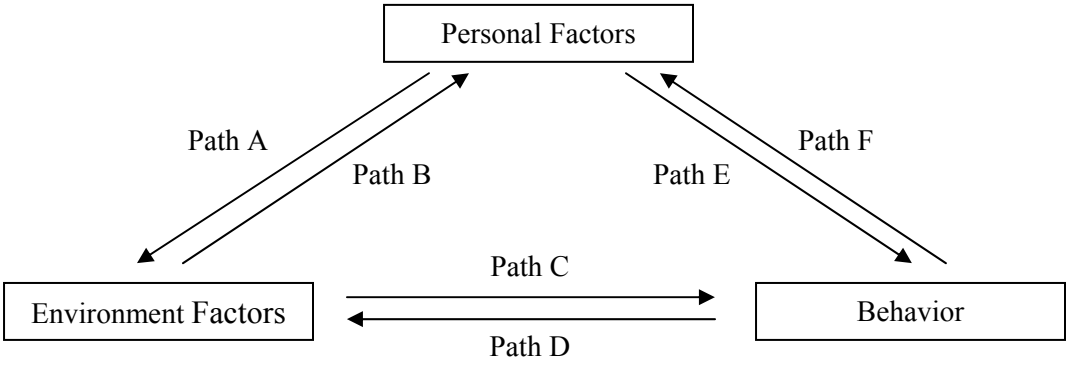
of loss or anger, imbalances in neurochemistry, or ineffective social skills, respectively (Bemporad, 1985; Clark et al., 1999; Lewinsohn, 1974). Other theories focus on intentions (Theory of Planned Behavior [Ajzen, 1991] or negative events (Learned Helplessness Theory [Abramson, Seligman, & Teasdale, 1978]). While these specialized theories may be preferred when investigating specific therapeutic factors or techniques, a theoretical framework explaining changes in psychosocial functioning during acute-phase treatment for MDD would best serve the field if it is comprehensive enough to apply to multiple treatment settings and strategies.

Second, social cognitive theory is parsimonious in it's incorporation of environmental, personal, and behavioral factors. While equally comprehensive, other theories considered for this dissertation were more complex, seeking to account for every unique interaction between factors (e.g., Luyten, Blatt, Houdenhove, & Corveleyn, 2006; O'Connor, 2003; Schotte, Bossche, Doncker, Claes, & Cosyns, 2006). While such complexity is at times beneficial for understanding complicated phenomena, a parsimonious theory is preferred when applying theory to a specific research question.

Third, social cognitive theory is inherently flexibility due to it's concept of triadic reciprocal causation (Bandura, 1986). Within this concept, environmental, personal, and behavioral factors interact with each other through bidirectional causal pathways (see Figure 1). In other words, according to social cognitive theory, behavior is not merely the result of environmental stimuli or altered maladaptive cognitions, but these factors work together to determine each other.

Figure 1

Model of Triadic Reciprocal Causation



Note. Adapted from Bandura (1986); Personal factors include emotions, cognitions, and biological processes; Environmental factors include life events, socioeconomic status, education, familial structures, etc.

Finally, social cognitive theory is backed by a body of literature that allows researchers to predict how factors interact under certain circumstances to produce behavior change. For instance, Bandura and his colleagues claim "economic conditions, socioeconomic status, and educational and family structures affect behavior largely through their impact on people's aspirations, sense of efficacy, personal standards, affective states, and other self-regulatory influences, rather than directly" (p. 15; Bandura, 2001). In other words, social cognitive theorists predict that the indirect paths, comprised of paths B and E, partially mediate the direct path between the environment and behavior (path C; see Figure 1). In so doing, social cognitive theory separates itself from models that are more descriptive in nature (e.g., Biopsychosocial Model, Diathesis-Stress Model, etc.) rather than explanatory.

What is Social Cognitive Theory?

First introduced by Miller and Dollard (1941) as social learning theory, social cognitive theory was created by researchers who rejected the idea held by strict behaviorists that behavior was simply the result of environmental stimuli. Instead, social cognitive theorists suggested that the environment's impact on human behavior was mediated by internal, cognitive mechanisms, such as observational learning (Bandura, 1962) and outcome expectancies (Mischel, 1973). Developing this concept further, Bandura (1986) claimed that intrapersonal factors worked together with the environment to determine behavior through a process called reciprocal determination (see Figure 1). Due to the centrality of triadic reciprocal causation to social cognitive theory, each component is described in greater detail.

Social cognitive theorists claim the environment is made up of factors that are physically external to the person, such as life events, culture, economic conditions, education, familial dynamics, etc. Bandura (2001) distinguished between three types of environments: imposed, selected, or constructed. These categorizations depend upon the degree to which a person works to create their environment, rather than simply react to it. For example, people experience an imposed environment when they do not control their surroundings, which can be a positive, neutral, or negative experience depending on how well the environment meets their needs. A selective environment is one in which people carry out certain behaviors to bring about desired outcomes. In a selected environment, people take part in creating their surroundings through their behavior, either eliciting favorable consequences or becoming trapped in patterns of negative reinforcement. Lastly, people experience a constructed environment when they play an active role in

creating their own social systems (e.g., economic condition, education, etc.), thereby increasing the amount of control they have over their lives.

According to social cognitive theory, personal factors are made up of biological, emotional, and cognitive influences. With regard to biological influences, social cognitive theory rejects a reductionist view that claims all human function can ultimately be reduced to biological determinants. Instead, biological processes are thought to merely set the stage for any number of human experiences that are ultimately determined by environmental pressures and a person's judgment on what is most adaptive. Whatever behaviors are chosen then reciprocally determine biological processes directly or through their impact on the environment. For example, human's use of tools led to improved living conditions, which ultimately resulted in improved survival rates and further biological advances.

Similarly, social cognitive theorists maintain that a person's emotions are reciprocally determined by other personal factors (e.g., cognitions and biological processes), behaviors, and environmental events (see Figure 1). For example, overly negative and self-defeating thoughts (e.g., Clark et al., 1999), dysregulation of the stress response (e.g., Gold & Chrousos, 2002), poor interpersonal skills (e.g., Lewinsohn, Hoberman, Teri, & Hautzinger, 1985), and adverse life events (e.g., Hammen, 2005; Oatley & Bolton, 1985) are all associated with depressed mood. Of course, the direction of this association works both ways, as depressed mood can also exacerbate these conditions.

While social cognitive theorists have attended to biological and emotional influences, most of their research centers on the role of cognitions. Cognitive concepts

such as self-regulation, self-efficacy, forethought, vicarious motivation, and behavioral capacity have all been successfully applied to bring about or better understand behavior change (Bandura, 1986, 1997). For example, researchers have applied social cognitive theory to improve health behaviors by: a) increasing people's confidence in performing target activities (self-efficacy), b) teaching negative consequences (outcome expectations), c) modeling and rewarding appropriate behavior (observational learning and vicarious motivation), or d) having people set personal goals for behavior change (self-regulation), to name a few (Baranowski, Perry, & Parcel, 1997). Also, if it could be substantiated that psychosocial treatments for depression have similar outcomes (see Lambert & Ogles, 2004), Bandura (1997) hypothesizes it would be because of the mediating effect of self efficacy, a cognitive construct defined as "people's judgments of their capabilities to organize and execute courses of action required to attain designated types of performances" (p. 391; Bandura, 1986). From this viewpoint, although psychosocial treatments employ different strategies, they may all work to improve depressive symptom severity and psychosocial functioning by improving self-efficacy (Cutrona & Troutman, 1986; Major et al., 1990).

In social cognitive theory, behavior is simply defined as a person's actions and categorized into three domains: personal, proxy, and collective (Bandura, 2001). People's behavior is considered personal when they alone carry out certain duties to achieve a desired outcome. Proxy behaviors occur when people think others are better able to secure desired outcomes, so they delegate responsibilities and rely on others to perform necessary actions. For example, when parents drop children off at daycare, they entrust their parenting responsibilities to other people and therefore fulfill their roles as

parents in proxy. Finally, collective behaviors occur when people work together as a group toward a common outcome. As explained in the next section, all three types of behavior can make up a person's psychosocial functioning if it fulfills a social role or obligation.

Conceptualization of Psychosocial Functioning Based on Social Cognitive Theory

Psychosocial functioning is defined as a person's performance in and satisfaction with "normal social roles", including occupational, interpersonal, and recreational domains (Hirschfeld, 2000). Within social cognitive theory and its model of triadic reciprocal causality (see Figure 1), psychosocial functioning represents a subset of behaviors carried out to fulfill social obligations. In other words, whereas the behavioral component in Figure 1 represents all human actions, psychosocial functioning represents only those behaviors carried out to meet normative social roles. To illustrate, gardening can be considered part of a person's psychosocial functioning, since it fulfills social roles pertaining to recreation, but tying a shoe cannot, since it does not fulfill social roles in occupational, interpersonal, or recreational domains.

That said, however, an absolute distinction cannot be made between behaviors that do or do not make up a person's psychosocial functioning. This is because the social roles that dictate a person's psychosocial functioning are unique to the individual and determined by his or her surroundings and set of standards. In this way, psychosocial functioning is reciprocally determined by environmental and personal factors. For example, while eating would not normally be considered part of someone's psychosocial functioning, food critics would include it because it fulfills a social role pertaining to

their occupation. What follows is a discussion of the reciprocal interactions between psychosocial functioning, the environment, and personal factors.

Psychosocial Functioning and Environmental Factors: In order to function in society, people need to know what is expected from them. Most of this instruction comes from people's surroundings in the form of normative social roles. In their book on psychosocial functioning in psychiatry, Tyrer & Casey (1993) maintain that psychosocial functioning is dependent on an understanding of a person's social status and class. Likewise, Clare, Corney, and Cairns (1984) claim it is essential for researchers to "locate the individual in his or her social context" (p. 326) in order to assess psychosocial functioning, which they define as the discrepancy between "actual behavior" and "ideal" social norms. Lastly, distinct social roles may exist across racial/ethnic groups. For instance, Weissman et al. (2001) report that ratings of psychosocial functioning, especially family functioning, differ significantly by race/ethnicity, with Hispanics and African Americans reporting "worse" functioning than non-Hispanic whites.

Psychosocial Functioning and Personal Factors: As mentioned before, psychosocial functioning is partly determined by the degree to which role fulfillment meets personally relevant standards, including both satisfaction with aspects of the role itself and the person's performance within it. So, feelings of satisfaction and beliefs about what can or should be accomplished (i.e., personal factors) play large roles in determining psychosocial functioning. Bandura (1997) suggested that poor psychosocial functioning most likely arose when people set standards for themselves that far exceeded their abilities. Likewise, Bosc et al. (1997) theorized that people with low self-efficacy beliefs were more likely to rate their interpersonal functioning as poor. Interestingly,

biological factors may also play a role, as neural pathways have been linked to interpersonal functioning and people's evaluation of it (e.g., Heim, Nemeroff, 2001; Meyer-Lindenberg et al., 2006).

Psychosocial Functioning, Environmental Factors, and Personal Factors: Social cognitive theorists predict personal factors mediate the impact of environmental factors on behavior. There is no reason to think this mediational relationship does not hold true when only considering behaviors that fulfill social roles (i.e., psychosocial functioning). In support of this assertion, Diener, Suh, Lucas, and Smith (1999) found factors such as income, race, and education effect a person's subjective evaluation of his or her work, family, and leisure activities.

How Does Social Cognitive Theory Account for Changes in Psychosocial Functioning during A-CT for MDD?

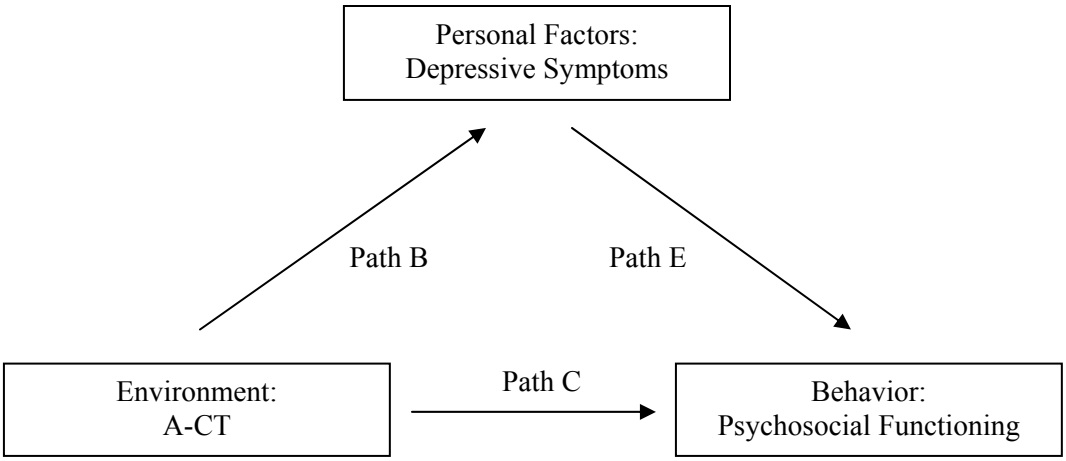
In this dissertation, an attempt is made to understand better how changes occur in psychosocial functioning during A-CT for MDD, both independently and in relation to changes in depressive symptom severity. The above conceptualization of psychosocial functioning with social cognitive theory provides a theoretical framework to help operationalize and interpret results from this research question. Since psychosocial functioning pertains to behaviors that fulfill social responsibilities, it can be operationalized as part of a person's behavior. Regarding depressive symptoms, diagnostic manuals (American Psychiatric Association, 2000) and confirmatory factor analyses (e.g., Maher, Mora, & Leventhal, 2006) report depressive symptoms consist of negative emotions (e.g., depressed mood, guilt), cognitions (e.g., diminished interest,

worthlessness, recurrent thoughts of death), and neurovegetative conditions (e.g., insomnia, fatigue, psychomotor retardation). These categories match those incorporated within the construct of personal factors (i.e., emotions, cognitions, and biological processes). Since depressive symptoms appear to represent a subset of emotions, cognitions, and biological processes, they are operationalized as personal factors.

To test one segment of reciprocal causation, one must determine the environmental context within which relationships between factors are examined. For this dissertation, the environmental context is A-CT. Taken together, this dissertation examines the segment of reciprocal causation between depressive symptoms and psychosocial functioning in an environment that is held constant and consists of A-CT (see Figure 2).

Figure 2

Operationalization of Primary Aim in Terms of Triadic Reciprocal Causation with Predicted Mediational Relationship



Note. Adapted from Bandura (1986); A-CT = Acute-Phase Cognitive Therapy; MDD = Major Depressive Disorder.

Given prior research (e.g., Hirschfeld et al., 2002; Lenderking et al., 1999; Vittengl et al., 2004), it is predicted that a) changes in depressive symptom severity partially mediate changes in psychosocial functioning and not vice versa (see Path E), and b) significant, independent changes occur in psychosocial functioning (see Path B) and depressive symptom severity (see Path C) during A-CT (see Figure 2). In the next section, potential implications of this finding are discussed.

Potential Implications for Social Cognitive Theory and Therapeutic Process

In summary, social cognitive theory provides a useful heuristic for better understanding how exposure to A-CT relates to change in psychosocial functioning. For

example, if the hypothesis is substantiated in which change in depressive symptom severity partially mediates change in psychosocial functioning, then change in personal factors partially mediates change in behaviors when individuals with MDD are exposed to A-CT (see Figure 2). In other words, the improvements in psychosocial functioning associated with A-CT may occur in individuals with MDD through changes in their maladaptive cognitions, negative emotions, and neurovegetative states.

If personal factors mediate changes in behavior when individuals with MDD are exposed to A-CT, it might also be concluded that internal processes require less time to change than behavior involving outside, environmental influences. To elaborated on this issue, Bandura (2001) claimed that most behavior involved the "coordination of interdependent plans of action" and the melding of "diverse self-interests in the service of common goals" (p. 7), a challenging pursuit at best. So, it could be said that behaviors designed to fulfill social roles depend upon the collaborative efforts of those involved (e.g., spouse, partner, parent, co-worker, teacher, etc.) (see paths C and D in Figure 1). It could also be the case that individuals with MDD require more time to change their psychosocial functioning because not only do they have to change, but other parties with whom they interact have to change. For example, in order for a depressed wife to improve her marital functioning, she has to change her manner of interacting with her husband, and her husband has to respond to and begin reciprocating this behavior. In sum, intrapersonal processes may take less time to change because they are less dependent upon these outside influences.

This study may also have practical implications. For instance, with an improved understanding of how psychosocial functioning changes during acute-phase treatment,

researchers may be better equipped to establish a consensus definition and standard technology to measure the construct. Said differently, people may have a better idea of what psychosocial functioning is or how to measure it if they know how it differs from, overlaps with, or is influenced by other elements of psychopathology. Also, with this knowledge researchers and clinicians may: a) better inform their clients suffering from recurrent MDD about when and how they are likely to experience improvements in their functioning during A-CT and b) more appropriately use instruments that measure psychosocial functioning when diagnosing or quantifying treatment effects.

How is Psychosocial Functioning Measured?

Over 30 instruments exist that measure psychosocial functioning (Goldman, Skodol, & Lave, 1992; Weissman et al., 1981; Tyrer et al., 1993; Wiersma, 1996), and the National Institute of Mental Health (NIMH) recommends that more be developed (NIMH, 2000). One reason for this proliferation of instruments despite the lack of a consensus definition may be that historically the investigation of psychosocial functioning has occurred secondarily to the study of mental illness. This tradition was reinforced in 1994, when impairments in psychosocial functioning were required in 130 axis I and axis II diagnoses (American Psychiatric Association, 1994), making functional impairment an integral part of many mental disorders (Spitzer et al., 1999). As a result, researchers have created specialized instruments of psychosocial functioning for individuals diagnosed with schizophrenia (e.g., Dickerson, Parente, & Ringel, 2000), major depressive disorder (e.g., Bosc, 2000), and personality disorders (e.g., Tyrer,

1990), to name a few. Similarly, instruments of psychosocial functioning have been tailored for inpatients (e.g., Baker & Hall, 1988), elderly populations (Branch & Jette, 1981), and couples experiencing marital distress (e.g., Kupfer & Frank, 1974; Spanier, 1976). Since these instruments were created without a consensus definition, they each assess different domains of psychosocial functioning that are sometimes unique to the disorder or population. While such specialization can be beneficial, it may also diversify the field and limit comparisons across studies, as most researchers and clinicians simply adopt a definition of psychosocial functioning that reflects how their instrument operationalizes the construct.

Given this, it would be beneficial to understand how extant instrumentation operationalizes the construct of psychosocial functioning. For example, instruments operationalize psychosocial functioning as one domain, like marital adjustment (e.g., Dyadic Adjustment Scale [DYS; Spanier, 1976]) or sexual functioning (e.g., Brief Sexual Function Questionnaire [BSFQ; Reynolds et al., 1988]), or multiple domains (e.g., Longitudinal Interval Follow-up Evaluation – Range of Impaired Functioning Tool [RIFT; Leon et al., 1999], Social Adjustment Scale-Self Report [SAS-SR; Weissman & Bothwell, 1976], etc.). Other instruments include components in their operationalization of psychosocial functioning that are related but not part of the construct, such as subjective feelings (e.g., Quality of Life Enjoyment and Satisfaction Questionnaire [Q-LES-Q; Endicott et al., 1993]) or physical health (Medical Outcome Study 36-item Short-Form Health Survey [SF-36; Ware & Sherbourne, 1992]). Most instruments try to separate role performance from environmental influences to avoid having to identify social norms for each individual (e.g., Social Adaptation Self-evaluation Scale [SASS;

Bosc et al., 1997]; Tyrer et al., 1993), while others seek to account for these factors (e.g., Social Maladjustment Schedule [SMS; Clare & Cairns, 1978]). Also, some instruments allow researchers and clinicians to combine estimates of symptom severity and psychosocial functioning into one, global scale of functioning (e.g., Global Assessment of Functioning [GAF; American Psychiatric Association, 2000]). While most instruments in this field are self-report (e.g., SAS-SR, DYS, etc.), some are clinician-rated (e.g., RIFT, GAF). Needless to say, existing instrumentation operationalizes psychosocial functioning differently (e.g., Weissman et al., 2001). Of note, researchers have yet to determine the extent to which these distinct conceptualizations represent converging paths that bring scientists to the same destination or divergent paths with diverse destinations.

In this study, psychosocial functioning is operationalized using the clinician-rated RIFT (Leon et al., 1999) and SAS-SR (Weissman & Bothwell, 1976). Total score from these measures are combined to create an index of psychosocial functioning, employing a procedure that has a proven track record (Vittengl, Clark, & Jarrett, 2004). The purpose for creating this index is to: a) more accurately assess the construct, as "scores across a set of measures tend to be more reliable and valid than scores on any individual measure" (p. 71; Kline, 2005), and b) control for shared method variance by including more than one source of data (i.e., client and evaluator). Also, as mentioned above, the creation of an index of psychosocial functioning is warranted given that existing instruments operationalize psychosocial functioning differently.

Since the RIFT has not been used to track changes in psychosocial functioning during A-CT and data with this instrument is collected retrospectively, steps are taken to

ensure its psychometric quality before including it in analyses. Based on its successful track record of measuring changes in psychosocial functioning among individuals diagnosed with MDD (e.g., Solomon et al., 2004; Vittengl et al., 2008), bipolar disorder (Judd et al., 2005), and Obsessive-Compulsive Disorder (Eisen et al., 2006), however, the RIFT is expected to have acceptable reliability and validity when compared to four other measures of psychosocial functioning examined by this study.

MDD Impairs Psychosocial Functioning by Definition and by the Data

As mentioned before, in order to be diagnosed with MDD, a person must experience "clinically significant distress" and/or functional impairment (American Psychiatric Association, 2000). Therefore, it logically follows that, by definition, MDD is associated with a high degree of impairment in psychosocial functioning. Since the early 1970's (e.g., Weissman, Paykel, Siegel, & Klerman, 1971), researchers have consistently validated this claim, showing that individuals diagnosed with MDD have greater impairments in all areas of psychosocial functioning, measured by the SAS-SR, SF-36 and Groningen Social Disability Schedule (GSDS; Wiersma, DeJong, & Ormel, 1988), than those without (Hays et al., 1995; Mintz, Mintz, Arruda, & Hwang, 1992; Ormel et al., 1993; Spijker et al., 2004). In fact, researchers using the SF-36 have found that the global impairments in psychosocial functioning associated with MDD rival those typically associated with chronic medical diseases (e.g., arthritis, cardiovascular disease, diabetes, cancer, etc.; Cassano & Fava, 2002; Hays et al.; Simon, 2003), and represent the "leading cause of disease-related disability among women in the world today" (p. 5;

Kessler, 2003). Impairments in psychosocial functioning, measured by the SF-36 and number of disability days, increase further when individuals with MDD are diagnosed with comorbid medical conditions (Klerman, 1989; Von Korff, Ormel, Katon, & Lin, 1992). Making matter worse, researchers report that psychosocial functioning appears relatively stable over time (Coryell et al., 1993; Judd, & Akiskal, 2000; Laroche, Hodgins, & Toupin, 1995; Ormel, Oldehinkel, Nolen, & Vollebergh, 2004), especially in the case of treatment-resistant MDD (Dunner et al., 2006) and MDD with psychotic features (Rothschild et al., 1993).

When examining the impact of MDD on psychosocial functioning with greater detail, researchers discover that impairments exist across each functional domain. For instance, in a series of studies following 40 women diagnosed with MDD, impairments in close interpersonal relationships, as measured by the SAS-SR, persisted over four-years, even after the complete remission of depressive symptoms (Bothwell & Weissman, 1977; Paykel et al., 1973; Weissman, et al 1971). As a consequence, these women were significantly less able to fulfill their roles as spouses than women who had never been depressed (Bothwell et al.). Similarly, in cross-sectional research, individuals with MDD reported having worse family functioning than individuals without MDD on the McMaster Family Assessment Device (Epstein, Baldwin, & Bishop, 1983), a measure that assesses seven areas of family functioning (e.g., communication, roles, general functioning, etc.; Saeki, Asukai, Miyake, Miguchi, & Yamawaki, 2002). Interestingly, family members of depressed patients also endorse greater impairments in psychosocial functioning than normative samples, suggesting that the family members might also be depressed or have suffered a loss in psychosocial functioning due to some yet to be

identified mechanism. In related research, immigrants with MDD showed more impairment on measures of acculturation stress and increased social isolation compared to immigrants without depression, showing how depression can negatively impact one's ability to interact with and adapt to their social surroundings (Haasen & Sardashti, 2000).

In addition, the vast majority of individuals with MDD endorsed occupational impairments, both in terms of lost productivity and absenteeism (Kessler, & Frank, 1997; Lepine, Gastpar, Mendlewicz, & Tylee, 1997; Marlowe, 2002; Stewart, Ricci, Chee, Hahn, & Morganstein, 2003; Tylee, Gastpar, Lepine, & Mendlewicz, 1999). In fact, Williams and Strasser (1999) found that depression cost employers more than any other behavioral disorder and accounted for more days missed at work than chronic medical conditions, such as diabetes, heart disease, hypertension, and lower back pain. While functional impairments in this area were alleviated with treatment that also effectively reduced depressive symptoms, changes took a considerable amount of time and were "undone" with each recurrence of a depressive episode (Mintz et al., 1992). Using the SAS-SR, De Lisio et al (1986) found this to also be the case with impairments in recreational functioning, which seemed to persist even when individuals were experiencing mild or no depressive symptomatology.

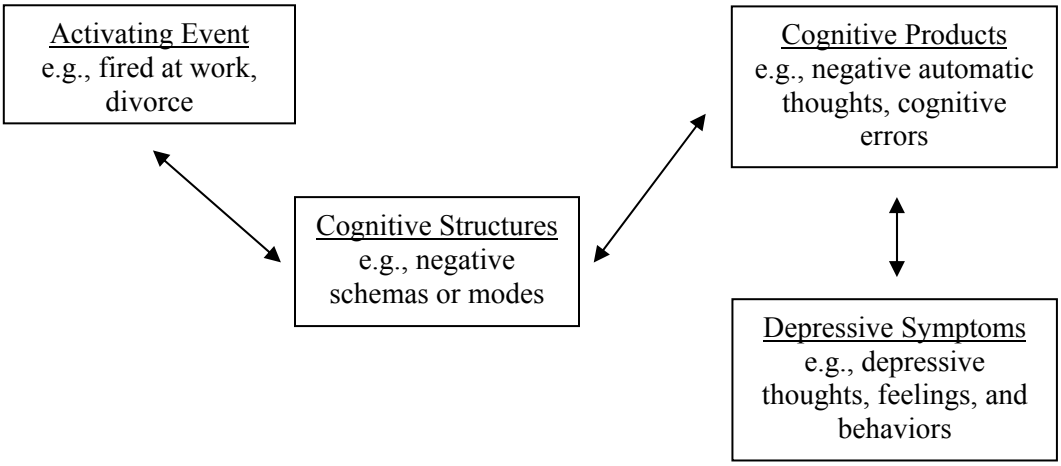
What is Cognitive Therapy?

According to the cognitive theory of depression, stressful internal or external events activate schemas (i.e., cognitive structures that help organize and provide meaning to the world; see Figure 3; Beck et al., 1979) that vary depending on their content,

including physiological, motivational, and behavioral schemas, to name a few (Clark et al., 1999). As seen in Figure 3, when activated, depressogenic schemas elicit patterns of negative thinking that lead to depressive thoughts, feelings, and behaviors (Beck et al.; Ingram, Miranda, & Segal, 1998). These depressive symptoms can also function to reinforce or perpetuate maladaptive thought patterns, forming a vicious cycle.

Figure 3

Cognitive Model of Depression



Note. This figure was adapted from a model proposed by Clark et al. (1999)

Cognitive therapy (CT; Beck et al., 1979) is a directive, time-specified therapeutic technique based on the cognitive theory of depression. In CT, therapists focus on changing maladaptive patterns of thought by helping depressed individuals identify, challenge, and ultimately replace them with more adaptive thoughts. In doing this, cognitive therapists teach patients about specific types of thought errors and help them identify and understand underlying schemas that perpetuate them. In addition,

cognitive therapists often teach patients how to monitor their thoughts, become more physically active, break up difficult tasks into ones that are more manageable, and try new behaviors to test old or new ways of thinking. By changing depressogenic schemas and the resultant maladaptive thought patterns, A-CT effectively reduces depressive thoughts, feelings, and somatic conditions (Beck et al., 1979; Butler et al., 2006; DeRubeis et al., 1990; Garratt et al., 2007; Hollon & Beck, 2004; Whisman, 1993).

Most Studies Suggest that Acute-Phase Cognitive Therapy Reduces Impairments in Psychosocial Functioning in Adults with MDD

Before explorations can be made into how psychosocial functioning changes during A-CT, it is first important to show that changes do indeed occur. A-CT is a promising psychosocial intervention that effectively reduces depressive symptoms (for a comprehensive meta-analysis, see Gloaguen, Cottraux, Cucherat, & Blackburn, 1998). However, less is known about its impact on psychosocial functioning (Hirschfeld et al., 2000). On one hand, in individuals with MDD, researchers suggest A-CT effectively reduces impairments in: a) overall psychosocial functioning, measured by the GAF (Elkin et al., 1989) and SAS-SR (Vittengl et al., 2004), b) marital functioning, measured by the DYS (Beach & O'Leary, 1986), c) family functioning, measured by a homemade rating mother-child relationship (Murray, Cooper, Wilson, & Romaniuk, 2003), d) satisfaction with role performance, measured by the Q-LES-Q satisfaction subscale (Saulsman, Coall, & Nathan, 2006), and e) occupational functioning, measured by the Work Potential Profile (Della-Posta & Drummond, 2006; Rowe, 1997) and SAS-SR work subscale

(Imber et al., 1990; Mintz et al., 1992). Researchers have replicated these findings in primary care, elderly, low income, and minority women populations, using the SAS-SR and SF-36 (e.g., Arean et al., 2005; Miranda et al., 2003; Thompson, Gallagher, & Breckenridge, 1987; Ward et al., 2000). Researchers have also found that A-CT reduces impairments in psychosocial functioning when provided on an individual, group, or computerized basis, using the Global Assessment Scale (GAS; Endicott, Spitzer, Fleiss, & Cohen, 1976), Work and Social Adjustment Scale (WSAS; Mundt, Marks, Shear, & Greist, 2002), and SAS-SR (e.g., Cavanagh et al., 2006; Gelhart & King, 2001; Imber et al., 1990; Proudfoot et al., 2004; Thase, Simons, Cahalane, & McGeary, 1991). On the other hand, researchers suggest A-CT without partner participation does not significantly reduce marital dysfunction, measured by the DYS (Beach & O'Leary, 1992) and Maudsley Marital Questionnaire (Crowe, 1978; Emanuels-Zuurveen & Emmelkamp, 1996, 1997), or sexual dysfunction, measured by the BSFQ (Nofzinger et al., 1993), in individuals with MDD. Also, no information exists in the literature on the effects of A-CT on impairments in recreational functioning in individuals with MDD.

Taken as a whole, most of the research indicates A-CT is effective at reducing impairments in psychosocial functioning in individuals with MDD. However, some questions remain regarding how effective A-CT is at treating the impairments in interpersonal and recreational functioning associated with MDD. That being said, it is unclear the extent to which differences in the literature are attributable to idiosyncrasies across study designs, patient populations, or measurement error, as different measures of psychosocial functioning were used in each study. It may also be the case that variations in the literature reflect differential treatment effects across domains of psychosocial

functioning (e.g., Miller et al., 1998; Weissman, Klerman, Prusoff, Hanson, & Paykel, 1976). For instance, impairments in interpersonal functioning may take longer to change than impairments at work during acute-phase treatment. Since it is important to further understand the impact of A-CT on psychosocial functioning, these issues were addressed in Chapter 3 through a quantitative review of the literature.

Impairment in Psychosocial Functioning Correlates Highly with Depressive Symptom Severity

A person's health is associated with their psychosocial functioning. This is especially the case with MDD, where certain aspects of a person's psychosocial functioning (e.g., social isolation, irritability around others, inability to go to work, etc.) may be considered symptomatic of underlying psychopathology (American Psychiatric Association, 2000). In fact, the high correlation between depressive symptom severity and psychosocial functioning may be one reason why a consensus definition and conceptual understanding of psychosocial functioning eludes researchers.

Using cross sectional data and a variety of instruments to assess each construct, researchers have found estimates of psychosocial functioning significantly correlate with depressive symptom severity (Clark, Vittengl, Dolores, & Jarrett, 2003; Finkelstein et al., 1996; Judd, Akiskal, et al., 2000; Lepine et al., 1997; Mulder, Joyce, & Frampton, 2003; Pedersen, Pallay, & Rudolph, 2002; Revicki, Turner, Brown, & Martindale, 1992; Sullivan, Adams, Thibault, Corbiere, & Stanish, 2006; Trivedi et al., 2006). Longitudinal data spanning three to 10 years substantiate this correlation, showing increases in

depressive symptom severity are associated with stepwise incremental increases in functioning impairments measured by the Longitudinal Interval Follow-up Evaluation – Psychosocial Interview (LIFE-PI; Keller et al., 1987) and GSDS (Judd, Akiskal, et al., 2000; Ormel et al., 1993; Ormel et al., 2004). When depressive symptoms are operationalized in terms of treatment response (i.e., ≥ 50% reduction in depressive symptoms), researchers report that treatment responders have better levels of psychosocial functioning on the GAS, SAS-SR, and SF-36 than non-responders at baseline, mid-treatment and post-treatment (Kennedy et al., 2001; Miller et al., 1998; Riso et al., 1997). In addition, impairments in psychosocial functioning, measured by the DYS, LIFE-PI, RIFT and SAS-SR, have been associated with relapse and recurrence (Hooley & Teasdale, 1989; Leon et al., 1999; Rodriguez et al., 2005; Solomon et al., 2004; Vittengl et al., 2008) and higher rates of residual depressive symptoms after acute-phase treatment (Ogrodniczuk, Piper, & Joyce, 2004).

The Relationship between Psychosocial Functioning and Depressive Symptom Severity needs further Clarification

Due to the high correlation between psychosocial functioning and depressive symptom severity, researchers have investigated the degree to which these constructs conceptually overlap. In addressing this issue from a diagnostic perspective, Spitzer et al (1999) reported that measures of psychosocial functioning and depressive symptom severity were "pragmatically redundant" because it was highly unlikely that an individual had a sufficient number of depressive symptoms to be diagnosed with MDD and not

experience clinically significant impairments in functioning (p. 1862). In support of this claim, Vittengl et al. (2004) reported "no significant change in social-interpersonal functioning independent of change in depressive symptoms" (p. 651) in a sample diagnosed with MDD receiving A-CT. In this analysis, Vittengl et al. operationalized psychosocial functioning, using scores from the SAS-SR and DYS, and depressive symptoms, using an index that combined total scores from the Beck Depression Inventory (BDI; Beck, Ward, Mendelson, Mock, & Erbaugh, 1961), Hamilton Rating Scale for Depression-17 (HRSD-17; Hamilton, 1960), and Inventory for Depressive Symptomatology-Self Report (IDS-SR; Rush et al., 1986). When psychosocial functioning was measured by the General Life Functioning Scale (GLF; Elkin, Parloff, Hadley, & Autry, 1985), Lenderking et al. (1999) also found no significant changes in psychosocial functioning after controlling for changes in depressive symptom severity that were measured by the Montgomery-Asberg Depression Rating Scale (MADRS; Montgomery & Asberg, 1979). On the other hand, when researchers only use the SAS-SR or GLF to measure psychosocial functioning and the HRSD-17 to measure depressive symptoms, significant changes in psychosocial functioning occurred independently from changes in depressive symptom severity (Finkelstein et al., 1996; Hirschfeld et al., 2002; Lenderking et al.: Vittengl et al.). Taken together, it is unclear whether these findings reflect independent treatment effects for psychosocial functioning that go beyond the effect of treatment on depressive symptom severity or differences across instruments.

In investigating the timing of change between psychosocial functioning and depressive symptoms, researchers have consistently reported that psychosocial functioning, especially interpersonal functioning, takes longer to change than depressive

symptom severity (Agosti & Stewart, 1998; Bothwell et al., 1977; Coryell et al., 1993; De Lisio et al., 1986; Giller, Bialos, Riddle, & Waldo, 1988; Mintz et al., 1992; Paykel et al., 1973; Paykel, Weissman, & Prusoff, 1978; Rounsaville, Prusoff, Weissman, 1980). However, naturalistic or cross-sectional designs that hinder estimations regarding causality limit most of these studies, and none address psychosocial functioning and depressive symptoms during A-CT. This study is the first to use structural equation modeling to examine the sequence and process by which psychosocial functioning changes, both independently and in conjunction with depressive symptom severity during A-CT for MDD.

Summary

Major Depressive Disorder is a highly prevalent and recurrent disorder that impairs an individual's ability to work, relate to others, and engage in leisurely activities. Surprisingly, researchers have not reached a consensus on how these impairments are best defined or measured, much less how or when they improve during A-CT. Consequently, researchers do not know how to interpret initial findings that suggest A-CT improves the impairment in psychosocial functioning experienced by individuals with MDD. This study will advance the field by: a) introducing a theoretical framework that conceptualizes changes in psychosocial functioning, b) using this framework to examine the relationship between changes in psychosocial functioning and depressive symptom severity in a sample of 470 individuals undergoing A-CT for MDD, and c) investigating the psychometric quality of measures used to assess psychosocial functioning.

As a result, this study will help the field understand how psychosocial functioning changes during A-CT in relation to depressive symptoms severity. With this knowledge, the field moves one step closer to understanding how at least one type of psychotherapy (i.e., A-CT) treats MDD and its associated functional impairment. Knowledge of how psychosocial functioning overlaps or is different from depressive symptom severity may also facilitate future efforts to explore its role in the development, course, and treatment of other disorders and disease. The extent to which these findings generalize to other psychotherapies or modalities of treatment is beyond the scope of this work.

CHAPTER THREE
Change in Psychosocial Functioning During A-CT: A Meta-Analytic Review

Rationale for Meta-Analysis

As reported earlier, research shows A-CT effectively reduces impairments in psychosocial functioning in individuals with MDD. However, questions remain regarding A-CT's effectiveness in treating impairments in interpersonal and recreation functioning, as well as how and when changes in psychosocial functioning occur with respect to changes in depressive symptom severity. Additionally, researchers have yet to explain discrepancies that exist in the field, which may reflect measurement error or actual differential treatment effects across study designs, patient populations, and/or domains of psychosocial functioning.

Some have tried to address these issues through qualitative reviews of the literature (e.g., Bech, 2005; Bosc, 2000; Hirschfeld et al., 2000; Kennedy et al., 2001; Weissman, 2000), but these qualitative efforts are limited by their inability to quantify effects without bias or determine the extent to which variations across studies are systematic or random (Bushman & Wells, 2001). Similarly, previous efforts to understand the relationship between changes in psychosocial functioning and depressive symptom severity produced mixed results (i.e., Hirschfeld et al., 2002; Lenderking et al., 1999; Vittengl et al., 2004). Consequently, a meta-analytic review was undertaken to respond to these issues and make recommendations for this and future investigations of psychosocial functioning and A-CT (Dunn, Vittengl, Jarrett, 2007).

Methodology for Meta-Analysis

A meta-analysis was carried out of studies investigating the effects of A-CT on the impairment in psychosocial functioning associated with MDD. Results clarify the extent to which: a) psychosocial functioning improves during A-CT, b) changes in psychosocial functioning during A-CT are accounted for by study or patient characteristics, c) changes in psychosocial functioning differ across treatments, and d) changes in psychosocial functioning account for changes in depressive symptoms or vice versa. In what follows, the methodology and results are presented, along with a discussion of implications in the context of existing research.

Identification of Research Studies

To find studies that investigated the impact of A-CT on psychosocial functioning in individuals diagnosed with MDD, multiple electronic databases were first searched (i.e., PsycINFO, MEDLINE, ERIC, Cochrane) for articles published through March 2007. Articles were selected that contained search terms, or key-word roots, about depression (e.g., affective disorder, mood disorder, unipolar, *depress*, major depressive disorder, or major depressive episode), psychosocial functioning (e.g., *psycho function*, social *function*, social *adjust*, social *adapt*, interpersonal *function*, marital discord, occupational *function*, *disabl*, work capacity, dysfunction, job status, employment status, leisure activity, recreation, or satisfaction), and cognitive-behavioral therapy (cognitive, *behavior*, cognitive therapy, cognitive behavioral therapy, or CBT). As a second step, reference sections of literature reviews or meta-analyses investigating related topics were

reviewed (e.g., Gloaguen et al., 1998; Haby, Donnelly, Corry, & Vos, 2006; Hirschfeld et al. 2000; Keller, 2003; Knight & Stre, 1999; Latimer & Sweet, 1984; Mintz et al., 1992). Articles were not excluded based on a restrictive definition of cognitive therapy, but all studies were considered that employed cognitive and/or behavioral interventions during acute-phase treatment. For example, included studies described their treatments as "cognitive therapy" or "cognitive behavioral therapy", or they primarily sought to change cognition and/or behavior in study participants.

To be included, articles had to: a) be published in a peer-reviewed journal, b) include only adults (i.e., ≥ 18 years of age) diagnosed with MDD, c) provide patients with an acute-phase intervention that addressed cognition and/or behavior as a primary component, and d) assess psychosocial functioning before and after treatment. Excluded articles: a) were not published in a peer reviewed journal, b) included individuals < 18 years of age, c) included individuals not diagnosed with MDD, d) used a study sample diagnosed with another primary mental illness (e.g., Schizophrenia, Schizoaffective disorder, Bipolar Disorder, Substance Dependence, etc.), e) used a study sample diagnosed with a primary general medical condition (e.g., Alzheimer's disease, Huntington's disease, cerebrovascular disease, cardiovascular disease, HIV, etc.), f) did not provide a primarily cognitive or behavioral intervention during the acute-phase, or g) did not provide data on pre- and post-treatment psychosocial functioning. A list of studies that initially appeared relevant but were later excluded appears in Appendix B, along with the reasons for their exclusion.

Variable and Effect Size Coding

Coded data included characteristics of the patients, treatment strategies, assessment protocols, and study designs (See Table 2). A common metric was used (e.g., *d*; Cohen, 1988) to present effect sizes, and only one effect size was extracted from each study to maintain statistical independence across analyses. Effect sizes for changes in psychosocial functioning during A-CT were calculated using the standardized mean gain, which is calculated using the following formula:

$$ES_{MG} = \frac{M_{T1} - M_{T2}}{s_p}$$

For this equation, mean total scores of psychosocial functioning at pre- and post-A-CT were subtracted and divided by their pooled standard deviation (Lipsey & Wilson, 2001). Positive effect sizes represented improvements in psychosocial functioning and vice versa.

When comparing the effects of A-CT on psychosocial functioning to other treatment conditions (i.e., antidepressant medication, other psychosocial interventions, and non-active controls), effect sizes were calculated using the following formula for the standardized mean difference:

$$ES_{MD} = \frac{M_{G1} - M_{G2}}{s_p}$$

For this equation, mean total scores of psychosocial functioning at post-treatment across the two treatment conditions were subtracted and divided by their pooled standard deviation (Lipsey et al., 2001). In this case, positive effect sizes favored cognitive therapy, and negative effect sizes favored the comparison treatment. To avoid hand-made mathematical errors, all effect sizes were calculated using an effect size calculator (Wilson, 2001).

Statistical Analyses used in the Meta-Analysis

To calculate overall effect sizes (ESG), each effect size was weighted by its inverse variance using the following equation:

$$ES_G = \frac{\Sigma (w_i \times ES_i)}{\Sigma w_i}$$

In this equation, w_i is the weighted inverse variance of each effect size and ES_i is each effect size where i is equal to 1 through k, with k equal to the number of effect sizes being averaged. Conventional norms were used (i.e., 0.80, 0.50, and 0.20 represent large, medium, and small effects, respectively [Cohen, 1988]) to examine overall effect size magnitude and 95% confidence intervals.

The Q statistic determined homogeneity of variances across effect sizes. If variances between effect sizes were found to be heterogeneous, which was determined by a significant Q statistic, random-effects weighted regressions and analyses of variance (ANOVA) determined the extent to which variables coded in this study accounted for a significant amount of variability in predicting effect sizes. Random-effects were chosen

rather than fixed-effects models because: a) the sample of studies did not represent the universe of studies, both past and future, on the topic, and b) all sources of variance between effect sizes, besides sampling error, were not known. Of note, random-effects models typically produce larger error estimates and wider confidence intervals than fixed-effect models. The random-effect modeling was based on maximum likelihood estimation, with an alpha level set at 0.05 for hypothesis testing. Lipsey et al.'s (2001) formulas and macros developed for SAS were used to determine the overall effect sizes and identify potential moderating variables.

A series of linear regressions determined the extent to which changes in depressive symptom severity mediated changes in psychosocial functioning during A-CT. In the first linear regression, change in psychosocial functioning, measured by the SAS-SR, predicted change in depressive symptom severity, measured by the HRSD-17. Specifically, difference scores that represented pre- to post-A-CT changes in HRSD-17 and SAS-SR total scores were entered as the criterion and predictor, respectively. A second linear regression analysis was then run after switching these two variables, such that change in depressive symptom severity predicted change in psychosocial functioning. As a last step, two additional linear regressions were run, substituting total scores on the BDI for total scores on the HRSD-17.

Description of Articles Included in the Meta-Analysis

Fourteen studies met search criteria (See Table 2). These studies provided outcome data for a total of 1056 participants, with an average sample size of 75 patients

per study. Patients were most likely White (91.05%), middle aged (*M* = 39.15), and female (73.60 %). Psychosocial functioning was measured using the Dyadic Adjustment Scale (DYS), Global Assessment Scale (GAS), Quality of Life Enjoyment and Satisfaction Questionnaire (Q-LES-Q), and Social Adjustment Scale-Self-Report (SAS-SR), and participants were treated with Marital Therapy, CBT, Imipramine, Interpersonal Therapy, Counseling, and Psychodynamic Therapy. Table 2 organized the studies according to the type of instrument used to measure psychosocial functioning.

Table 2

Studies Included in Meta-Analysis of Changes in Psychosocial Functioning After Exposure to A-CT

Study	Definition of MDD	# of Weeks	# of Sessions	Psychosocial Functioning Measure	Psychosocial Functioning Domain	Design	Sample Size	Treatment Type	Effect Size
Beach et al 1986	MDD & DYS < 115	No info	14	DYS	Interpersonal functioning	Randomized Controlled Trail [a]	3	CBT	1.6208
							3	MT	1.7111
Jacobson et al 1991	MDD, HRSD ≥ 14, & BDI ≥ 20	No info	20	DYS	Interpersonal functioning	Randomized Trial, no control	20	CBT	0.3863
							19	MT	0.5260
							21	CBT+ MT	0.7486
Hollon et al 1992	MDD, HRSD ≥ 14, & BDI ≥ 20	12	20	GAS	Overall Psychosocial Functioning	Randomized Trial, no control	16	CBT	2.4239
							31	MED	2.6621
							16	CBT+ MED	3.0889
Thase, Bowler et al 1991	MDD & HRSD ≥ 15	4	20	GAS	Overall Psychosocial Functioning	Single Group Pre-post	16	CBT	3.5287

Study	Definition of MDD	# of Weeks	# of Sessions	Psychosocial Functioning Measure	Psychosocial Functioning Domain	Design	Sample Size	Treatment Type	Effect Size
Thase et al 1994	MDD & HRSD ≥ 14	16	26	GAS	Overall Psychosocial Functioning	Single Group Pre-post	45	CBT	2.0836
Thase, Simons et al 1991	MDD & HRSD ≥ 15	16	20	GAS	Overall Psychosocial Functioning	Single Group Pre-post	59	CBT	2.8340
Murray et al 2003	MDD	10	10	Homemade [b]	Interpersonal functioning	Randomized Controlled Trial	41	CBT	1.0911
							40	COUN	0.4443
							40	PSYD	0.7777
							31	Control	0.3624
Saulsman et al 2006	MDD	10	10	Q-LES-Q (Overall Satisfaction Subscale)	Satisfaction	Single Group Pre-post	109	CBT	0.7683

Study	Definition of MDD	# of Weeks	# of Sessions	Psychosocial Functioning Measure	Domain	Design	Sample Size	Treatment Type	Effect Size
Imber et al 1990	MDD & HRSD ≥ 14	16	20	SAS-SR [c]	Overall Psychosocial Functioning	Randomized Controlled Trial	37	CBT	1.5291
							47	IPT	1.3735
							36	MED	1.5385
							33	Control	1.2902
Propst et al 1992	MDD & HRSD-28 ≥ 14	13	18	SAS-SR	Overall Psychosocial Functioning	Randomized Controlled Trial	19	CBT	1.1572
							10	COUN	0.6696
							11	Control	0.3170
Rehm et al 1987	MDD, BDI ≥ 20, & MMPI-Depression Clinical Scale ≥ 70	10	No info	SAS-SR	Overall Psychosocial Functioning	Single Group Pre-post	32	CBT	1.4191
Shapiro et al 1994	MDD & BDI ≥ 16	8 or 16	8 or 16	SAS-SR (Social & Leisure Subscale)	Interpersonal functioning	Randomized Trial, no control	58	CBT	1.1619
							57	IPT+ PSYD	1.0202

Study	Definition of MDD	# of Weeks	# of Sessions	Psychosocial Functioning Measure	Psychosocial Functioning Domain	Design	Sample Size	Treatment Type	Effect Size
Thompson et al 1987	MDD, HRSD ≥ 14, & BDI ≥ 17	16	20	Multiple measures [d]	Overall Psychosocial Functioning	Randomized Controlled Trail [a]	26	CBT	1.1050 [e]
							24	PSYD	0.9917 [e]
Vittengl et al 2004	MDD & HRSD ≥ 16	14	20	Multiple measures [f]	Overall Psychosocial Functioning	Single Group Pre-post	156	CBT	0.8300

Note. MDD = Major Depressive Disorder; BDI = Beck Depression Inventory; MT = Marital Therapy; CBT = Cognitive Behavioral Therapy; COUN = Counseling; DYS = Dyadic Adjustment Scale; GAS = Global Assessment Scale; HRSD = Hamilton Rating Scale for Depression; IPT = Interpersonal Therapy; MED = Antidepressant Treatment; PSYD = Psychodynamic Therapy; Q-LES-Q = Quality of Life Enjoyment and Satisfaction Questionnaire; SAS-SR = Social Adjustment Scale-Self-Report. [a] no pre- to post-treatment scores of psychosocial functioning were provided for the control group; [b] Homemade instrument measured mother-infant relationship problems; [c] Pooled across the Social and Leisure and Work Subscales of the SAS-SR; [d] Pooled across the SAS-SR and GAS; [e] Pooled across Cognitive Therapy and Behavioral Therapy groups; [f] Pooled across the SAS-SR and DYS.

To What Extent Does Psychosocial Functioning Improve After Exposure to A-CT Compared to Pre-Treatment Baseline in Individuals Diagnosed with MDD?

To answer this research question, weighted, standardized mean gain effect sizes were calculated, using random-effects modeling based on maximum likelihood estimation. Across 14 studies that assessed pre-post changes in psychosocial functioning during A-CT, the random-effects weighted average effect size equaled 1.39 (SE = 0.13, p < 0.00001, 95% confidence interval = 1.13, 1.65). This finding shows A-CT effectively reduces the overall impairments in psychosocial functioning experienced by people with MDD, with an average effect size exceeding Cohen's (1988) recommendation for a large effect (e.g., d = 0.80). Effect sizes ranged from 0.39 to 3.53, and the index of heterogeneity across studies was statistically significant, $Q(13) = 193.79$, $p < 0.00001$. Therefore, additional analyses searched for potential moderators to account for the variability across effect sizes.

Does Instrument Type Accounts for a Significant Amount of Variance in Predicting Improvements in Psychosocial Functioning After A-CT?

To answer this question, a random-effects weighted ANOVA was run based on maximum likelihood estimation. In this analysis, the type of instrument was the independent variable, and the weighted effect sizes were the dependent variable. As seen in Table 3, information was provided for five different instruments of psychosocial functioning. A sixth variable was created, called "multiple measures," by combining

information across multiple instruments in two studies (i.e., Thompson et al., 1987, Vittengl et al., 2004). As a result, these two studies were represented by one effect size and not unduly weighted in the overall analysis by having more effect sizes than other studies.

The type of instrument used to measure psychosocial functioning accounted for 86.13% of the variance in predicting effect sizes, $Q(5) = 62.51$, $p < 0.0001$. In other words, the type of instrument played a significant role in determining the extent to which psychosocial functioning improved after exposure to A-CT. As seen in Table 3, the highest weighted, average effect size (i.e., 2.52) was associated with the GAS, which was nearly double that of the next highest effect size. This finding suggests that combining estimates of symptom severity and psychosocial functioning may improve the sensitivity of the measure to changes during treatment. The DYS had the lowest weighted, average effect size (i.e., 0.50), which supports previous research showing marital functioning is less likely to change as a result of A-CT without partner participation. The other instruments had moderately large to large weighted, average effect sizes ranging from 0.77 to 1.31, respectively (see Table 3). Of note, the GAS and SAS-SR were the most often used instruments of psychosocial functioning.

Table 3

Weighted Effect Sizes and 95% Confidence Intervals across Instruments Used to Assess Psychosocial Functioning during A-CT

Type of Measurement	*n*	Avg. ES	Standard Error	-95% CI	+95% CI	*t*-test	*p*
Dyadic Adjustment Scale	2	0.50	0.22	0.08	0.93	2.32	0.05
Global Assessment Scale	4	2.52	0.20	2.14	2.91	12.83	< 0.01
Homemade Instrument	1	1.09	0.25	0.61	1.58	4.42	< 0.01
Quality of Life Enjoyment and Satisfaction Questionnaire [a]	1	0.77	0.22	0.34	1.20	3.49	< 0.01
Social Adjustment Scale-Self-Report [b]	4	1.31	0.14	1.04	1.57	9.68	< 0.01
Multiple Measures	2	0.94	0.17	0.61	1.27	5.57	< 0.01

[a] Effect size only represents the General Satisfaction Subscale. [b] Includes two effect sizes that are based on subscales rather than the total score.

Does the Domain of Functioning Account for a Significant Amount of Variance in Predicting Improvements in Psychosocial Functioning After A-CT?

A random-effects weighted ANOVA based on maximum likelihood estimation answered this research question. In this analysis, the domain of psychosocial functioning measured by each instrument made up the independent variable, and the dependent variable consisted of weighted effect sizes. Nine studies that employed the GAS and/or SAS-SR were categorized as measuring overall psychosocial functioning. Four studies,

employing the DYS or homemade assessments of mother-infant interaction, measured interpersonal functioning, and only one study measured satisfaction. Research by Vittengl et al. (2004), which employed the SAS-SR and DYS, was categorized as measuring interpersonal functioning because the SAS-SR placed heavy emphasis on interpersonal functioning and the study's overall effect size was substantially lowered by the inclusion of the DYS.

The domain of psychosocial functioning did not account for a significant proportion of variance in predicting effect sizes, $Q(2) = 5.44$, $p < 0.07$. In other words, the domain of psychosocial functioning being assessed did not play a significant role in determining the extent to which psychosocial functioning improved after exposure to A-CT. However, the effect size for studies measuring overall psychosocial functioning were nearly double those measuring interpersonal functioning and satisfaction. Specifically, the weighted, average effect sizes for studies measuring overall psychosocial functioning, interpersonal functioning, and satisfaction were 1.73, 0.97, and 0.77, respectively.

Does the Study Design or Length of Treatment Moderate Changes in Psychosocial Functioning After A-CT?

For these analyses, study design served as the independent variable in a random-effects weighted ANOVA, and the length of treatment served as the predictor in a random-effects weighted regression. Weighted effect sizes made up the dependent variable, and both analyses used maximum likelihood estimation. Studies were

categorized as having a randomized-controlled (five studies), randomized (three studies), non-randomized (one study), or single group pre-post (five studies) design. Also, treatment length was coded by entering the number of protocol sessions and weeks of A-CT for each study. Treatment protocols ranged from 10 to 26 sessions and four to 16 weeks, averaging 18 sessions and 12 weeks, respectively.

Treatment design (i.e.,) did not account for a significant proportion of variance in predicting effect sizes, $Q(3) = 2.03$, $p < 0.57$. Furthermore, the number of sessions (Q (1) = 2.59, $p < 0.11$) or weeks of protocol treatment (Q (1) = 0.26, $p < 0.61$) did not account for a significant proportion of effect size variability. In other words, the quality of study design or length of treatment did not significantly influence changes in psychosocial functioning during A-CT.

Do Individuals with MDD Assigned to A-CT have Greater Improvements in Psychosocial Functioning than Those Assigned to Other Treatments or Control Conditions?

This research question was addressed with three separate analyses. For each analysis, weighted, standardized mean difference effect sizes were calculated, using random-effects modeling based on maximum likelihood estimation. In the first analysis, A-CT improved impairments in psychosocial functioning significantly more than non-active controls, with $d = 0.38$ ($CI_{95} = 0.07$ to 0.69; $p = 0.02$). This was a small to medium effect according to conventional norms (e.g., Cohen, 1988), and only three studies contributed data to this analysis (Imber et al., 1990; Murray et al., 2003; Propst, Ostrom,

Watkins, Dean, & Mashburn, 1992). Seven (Beach et al., 1986; Imber et al.; Jacobson, Dobson, Fruzzetti, Schmaling, & Salusky, 1991; Murray et al.; Propst et al.; Thompson et al., 1987; Shapiro et al., 1994) and two (Hollon et al., 1992; Imber et al.) studies contributed data to the second and third analyses, respectively. In these two analyses, A-CT did not differ from other psychosocial interventions (d = 0.11; CI_{95} = -0.04 to 0.26; p = 0.16) or antidepressant medication (d = -0.03; CI_{95} = -0.40 to 0.34; p = 0.88) in improving psychosocial functioning in individuals with MDD. Effect sizes did not vary significantly across all three analyses.

In drawing conclusions from these comparative analyses, A-CT appeared to improve impairments in psychosocial functioning associated with MDD more than non-active controls. At this time, however, one could not conclude that the effect of A-CT was different from other pharmacological or psychosocial interventions (e.g., generic counseling and Interpersonal, Marital, and Psychodynamic Therapies) on psychosocial functioning. Given the small number of studies reporting data, these conclusions may change as new study findings are taken into account.

Do Improvements in Psychosocial Functioning After A-CT Occur Independently of Changes in Depressive Symptoms?

When measured with the HRSD-17 (R^2 = 0.36, p = 0.0002) or BDI (R^2 = 0.20, p = 0.005), depressive symptom severity and psychosocial functioning shared a significant proportion of variability in pre- to post-A-CT change. Regarding independence of change, depressive symptoms were found to change significantly (i.e., HRSD: t = 17.24,

$p < 0.0001$; BDI: $t = 15.11$, $p < 0.0001$) when the variance associated with changes in psychosocial functioning was controlled. In other words, it appeared that improvements in depressive symptoms occurred independently of improvements in psychosocial functioning.

Results concerning the independence of changes in psychosocial functioning varied. On one hand, when controlling for the variance associated with changes in HRSD-17 total scores, psychosocial functioning changed significantly ($t = -2.46$, $p = 0.02$). From this result, one could conclude that psychosocial functioning changed independently of changes in depressive symptoms during A-CT. On the other hand, when controlling for the variance associated with changes in BDI total scores, psychosocial functioning did not change significantly ($t = -1.22$, $p = 0.23$). In this case, changes in depressive symptom severity appeared to fully account for changes in psychosocial functioning during A-CT.

To explain these conflicting results, Vittengl et al. (2004) speculated that the HRSD-17 "taps aspects of depressive symptoms less overlapping with social adjustment than other commonly used depressive symptoms measures" (p. 653-654), like the BDI. Consequently, the answer to the question, "Does psychosocial functioning change independently from changes in depressive symptoms during A-CT?" may depend on how the variables are measured. For instance, the answer would be *yes* if psychosocial functioning is measured with the SAS-SR and depressive symptom severity is measured with the HRSD-17. On the other hand, the answer would be *no* if psychosocial functioning is measured with the SAS-SR and depressive symptom severity is measured with the BDI.

Conclusions and Implications for Additional Research

This meta-analysis of 14 studies, including 1056 adults with MDD, suggested that A-CT substantially reduced the impairments in psychosocial functioning experienced by individuals with MDD (d = 1.39). Comparative analyses showed that A-CT improved psychosocial functioning significantly more than non-active controls but not more than other psychosocial interventions (e.g., generic counseling and Interpersonal, Marital, and Psychodynamic Therapies) or antidepressant medication. A high degree of variability between effect sizes existed in this data set, indicating the presence of moderating influences. The type of instrument used to measure psychosocial functioning accounted for the majority of this variability (86%), suggesting instruments of psychosocial functioning do indeed assess distinct constructs (Weissman et al., 2001). In other words, the effectiveness of A-CT in reducing the functional impairments associated with MDD appeared to depend on how the construct of psychosocial functioning was measured. For instance, when using the GAS, psychosocial functioning improved approximately three standard deviations from pre- to post-A-CT. However, when using the DYS, psychosocial functioning only improved half a standard deviation from pre- to post-A-CT.

These meta-analytic results also suggested the domain of psychosocial functioning played a role in how the construct changed during A-CT, with interpersonal functioning changing the least. While not statistically significant, this finding highlighted

the need for additional research to clarify the issue that controlled for the type of instrument used and included more domains of psychosocial functioning.

Additional analyses were inconclusive regarding the extent to which changes in psychosocial functioning occurred independently of changes in depressive symptom severity. However, results substantiated past research (i.e., Hirschfeld et al., 2002; Vittengl et al., 2004) suggesting changes in psychosocial functioning, measured by the SAS-SR, were only independent of changes in HRSD-17 total scores, a finding that may reflect measurement issues more than independent treatment effects.

Taken together, the above meta-analysis supported the use of A-CT in treating the occupational, interpersonal, and recreational impairments experienced by individuals with MDD, thereby alleviating the "real world" burden of this disorder. Researchers need to: a) replicate these findings using a wider array of patient populations, measures, and treatment strategies, b) investigate the degree to which instruments differ in measuring psychosocial functioning, c) determine how the construct is best defined and measured, and d) clarify the relationship between changes in psychosocial functioning and depressive symptom severity. The current study seeks to address these needs by:

1) Replicating the finding that suggests A-CT reduces impairments in psychosocial functioning with the largest known sample of individuals receiving A-CT for MDD,
2) Introducing a theoretical framework with which changes in psychosocial functioning can be conceptualized,

3) Comparing the psychometric quality of five measures of psychosocial functioning to justify inclusion of RIFT and SAS-SR data on the index of psychosocial functioning,
4) And applying structural equation modeling to clarify how psychosocial functioning changes during A-CT, both independently and in relation to depressive symptom severity.

As a result, this dissertation advances the field's knowledge of what psychosocial functioning is and how it changes during A-CT. This information could further efforts to develop a consensus definition and standard technology to assess psychosocial functioning and increase the field's knowledge of the role of the construct in diagnosis and treatment evaluation.

CHAPTER FOUR
Methodology

Overview of Methodology for Dissertation

This dissertation analyzed data from an ongoing, two-site, randomized controlled trial (RCT) sponsored by the National Institute of Mental Health (NIMH), entitled "Prophylactic Cognitive Therapy of Depression" (R01 MH 58397 and 58396) and "Are Cognitive Therapy's Antidepressant Effects Durable?" (R01 MH 69619 and 69618). Principal investigators Robin B. Jarrett, Ph.D., and Michael E. Thase, M.D., conducted this three-phase investigation at The University of Texas Southwestern Medical Center at Dallas (UT Southwestern) and Western Psychiatric Institute and Clinic at Pittsburgh (WPIC).

In the first, or acute phase, consented adult outpatients who were diagnosed with recurrent MDD and had a HRSD-17 $\geq$ 14 received 16 to 20 sessions of A-CT. Study patients received sessions twice weekly for four weeks, at which time they were categorized as early or late responders. If patients were early responders (i.e., $\geq$ 40% reduction in HRSD-17 total score by the eighth session of A-CT compared to diagnostic follow-up), they received sessions once a week for the final eight weeks. If they were late responders (i.e., < 40% reduction in HRSD-17 total score by the eighth session of A-CT compared to diagnostic follow-up), they continued receiving CT twice weekly for another 4 weeks and then one session a week for the final four weeks of the acute-phase.

At the end of A-CT, patients were stratified into lower or higher risk groups. Individuals in the higher-risk group who did not meet criteria for MDD and had a final HRSD-17 of ≤ 12 entered the second, or continuation phase of the study. In the continuation phase, patients were randomized to receive continuation-CT, Fluoxetine (Prozac), or pill placebo. After eight months of continuation treatment, all protocol treatments were discontinued, and patients entered the third and final phase of the study, a two-year follow-up.

Patients

For this study, research teams at UT Southwestern and WPIC entered patients' between the years 2000 to 2007 (see Appendix C for a description of this investigator's participation in data collection). As this study was conducted at two sites, demographic and clinical data were presented for the whole group and for each site. By the August 10, 2007, cutoff, a total of 484 patients met inclusion criteria and consented to enter A-CT (251 at UT Southwestern and 233 at WPIC). Of these, 470 patients were no longer active in the acute-phase and were included in study analyses.

Recruitment

The majority of patients were recruited using newspaper, Internet, and bulletin board announcements. Research teams at each site also informed community mental health centers and medical clinics of the project, encouraging referrals from mental health practitioners and primary care physicians. To ensure adequate representation within the

study of all groups, recruitment efforts extended to churches and community centers in African-American and English-speaking Hispanic communities. Through these efforts, the project was promoted widely throughout both Dallas and Pittsburgh.

Informed Consent

Research teams at UT Southwestern and WPIC provided potential patients information about the study during the informed consent process, including the risks, benefits, and alternatives to participation. As a result, potential patients were informed: a) that participation in the study was voluntary, b) they could withdrawal from the study at any time, c) about study expectations, and d) about additional treatment options. A data safety and monitoring board (DSMB), comprised of individuals uninvolved in the research, met annually and continually determined that the benefits of study participation outweighed the potential risks.

This informed consent process involved two steps. As a first step, interested patients signed an informed consent (see consent form for diagnostic evaluation in Appendix D) to undergo a diagnostic evaluation where information was gathered about their current and past symptoms and medical and social history. The two primary goals of this initial evaluation were to determine if patients met inclusion criteria and whether study treatment was appropriate. If the evaluator or patient deemed alternative treatment(s) were more beneficial, evaluators helped patients locate them. On the other hand, if the evaluator and patient considered study treatment appropriate and the patient met eligibility criteria, the evaluator scheduled the patient for a diagnostic follow up evaluation at least one week later.

The goals for the diagnostic follow up evaluation were to reassess eligibility criteria, review the previous evaluation, and again make recommendations about appropriate treatment. If the follow-up evaluator confirmed the patients' eligibility and appropriateness of study treatment, patients viewed a videotape that reviewed the consent process and study procedures. After viewing the videotape, patients could enter the study by signing a second informed consent (see consent form to enter A-CT in Appendix E). Patients renewed their consent after the acute and continuation phases in order to continue participation in the study.

Inclusion/Exclusion Criteria

To be included in the study, patients: a) met DSM-IV criteria for current MDE, b) scored ≥ 14 on the HRSD-17 at both the diagnostic evaluation and follow-up, c) had at least two episodes of MDD in lifetime, including current episode, d) provided written informed consent, e) spoke and read English, f) had an interest in psychotherapy, cognitive therapy, and research, and g) were aged 18-70 years. Diagnoses were made with the Structured Clinical Interview for DSM-IV (SCID-I; First, Spitzer, Gibbon, & Williams, 1996).

Patients were excluded if they: a) had severe or poorly controlled concurrent medical disorders or took medication that could cause changes in mood; b) could not provide informed consent; c) suffered from any psychotic or organic mental disorder, bipolar disorder, active alcohol or drug dependence, primary obsessive compulsive disorder, or primary eating disorders, with primary referring to the disorder associated with the most impairment or distress based on the evaluator's judgment; d) scored less

than 14 on the HRSD-17 at either the initial or a second diagnostic interview conducted at least 5 days later; e) could not complete questionnaires written in English; f) represented an active suicide risk requiring hospitalization or urgent care; g) had previously failed to respond to a trial of at least 8 weeks of CT conducted by a certified therapist; h) had previously failed to respond to at least 6 weeks of 40 mg of fluoxetine; or i) were pregnant or planned to become pregnant during the first 11 months after intake. The last exclusionary criterion was used because study patients could potentially receive continuation treatment with fluoxetine after A-CT. As needed, evaluators consulted the study physician regarding medical histories and, if indicated, required a physical examination and/or appropriate laboratory tests to ensure patient eligibility.

Therapist Competence and Adherence to Cognitive Therapy Model

To assure therapist competence and adherence, each of the 15 therapists in this study completed extensive training in cognitive therapy and achieved and maintained Cognitive Therapy Scale (CTS; Young & Beck, 1980) scores ≥ 40 before treating study patients. To maintain the high quality of CT, weekly group supervision occurred at each site. At supervision, therapists regularly (i.e., approximately four times a year) submitted videotaped sessions for review. Group supervisors (i.e., Robin B. Jarrett, PhD, at UT Southwestern and Sandra Kornblith, PhD, at WPIC) and other therapists observed and rated these videotaped sessions on the CTS, providing the target therapist with feedback on their strengths and weaknesses. Supervisors provided additional instruction if: a) at any time the supervisor judged the therapist to have difficulty applying the cognitive

model, b) the therapist requested consultation, or c) supervision resulted in CTS scores below 40. Response and attrition rates, as well as patient-reported alliance, were also monitored to track therapist competence in providing CT.

Study Procedure

Evaluators worked with prescribing physicians and patients to facilitate withdrawal from psychotropic medication at least one week prior to study entry, or more depending on the half-life of the medication and consultation by subspecialty. Also, patients agreed to defer additional psychotropic or pharmacological treatment for depression other than those provided by the study until relapse or the end of the study. Those who met inclusion criteria and signed consent received 16 to 20 sessions of A-CT. After completion of the acute-phase, independent evaluators blinded to treatment conditions examined patients (see Blind Evaluation on p. 78). Table 4 outlines the steps involved in the acute-phase of the study.

Table 4

Outline of Study Procedures for Diagnostic and Acute Phases, by Month, Week, and Event

Procedure	Months in Study	Week	Event		
Diagnostic Evaluation and Diagnostic Follow-up	0	-1	Triage evaluation, psychiatric evaluation, medical screen and drug wash-out, informed consent, 1st Psychoeducational Session		
			HRSD-17 reduction from diagnostic follow-up to week 4 of A-CT		
			≥40% (early responders)	<40% (late responders)	
Acute-Phase Cognitive Therapy	1	1	1-2	1-2	
		2	3-4	3-4	
		3	5-6	5-6	
		4	7-8	7-8	
	2	5	9	9-10	
		6	10	11-12	2nd Psycho-
		7	11	13-14	educational
		8	12	15-16	Session at
	3	9	13	17	session 11
		10	14	18	
		11	15	19	
		12-14 [a]	16	20	
Blind Evaluation	4	13-15			

Note. Blind Evaluations occur within seven days of the end of the acute-phase. [a] Two extra weeks allow time to make up missed sessions.

Diagnostic Evaluation (Visit 002) and Diagnostic Follow-up (Visit 003)

Study patients' first visit at each research site was the diagnostic evaluation (visit 002). At this visit, experienced evaluators obtained written consent for the diagnostic evaluation, determined each patient's lifetime psychiatric diagnoses, and verified

eligibility given the above inclusion/exclusion criteria. For these purposes, evaluators administered the HRSD-17, SCID-I, and an extensive medical and social history over the course of two to five hours. If patients met study criteria, evaluators scheduled them for a follow-up diagnostic evaluation one week later with a licensed, faculty-level clinician.

During the follow-up diagnostic evaluation (visit 003), evaluators determined whether there was any significant change in symptoms or patient history, reviewed diagnoses, and confirmed that the patient met study criteria, paying extra attention to questions posed by the initial evaluator. To do this, follow-up evaluators again administered the HRSD-17, reviewed the SCID-I from the diagnostic evaluation, checked the nature and quantity of medication consumed since the previous evaluation, and asked the patient to clarify any changes between evaluations. Initial and follow-up evaluators sought to reach a consensus on patients' diagnoses and dispositions, considering all available data. While disagreements were recorded and discussed, the follow-up evaluator made the final diagnoses and disposition. If patients met eligibility criteria and study treatment was deemed appropriate, the follow-up evaluator had each patient sign a written consent to enter the study. Once this consent was signed, they scheduled the patients' first session of A-CT to fall within five days of the follow-up evaluation. At both evaluations, if the patient was not eligible or it was determined that the study treatment was not appropriate, evaluators recommended and assisted patients in pursuing alternative treatment plans.

In carrying out the above diagnostic procedures, the two research clinics in this study (i.e., UT Southwestern and WPIC) differed slightly. For instance, because WPIC operated in a hospital setting, patients referred for study participation first underwent a

required, structured psychiatric evaluation by hospital staff. If they met inclusion/exclusion criteria, they were then scheduled for an initial diagnostic evaluation (visit 002). Patients at UT Southwestern did not receive a structure psychiatric evaluation before their initial diagnostic evaluation (visit 002). Patients at both sites were then scheduled for a diagnostic follow-up (visit 003) that occurred at least one week after the diagnostic evaluation and five days before beginning A-CT.

Evaluators at both sites correctly implemented the study's inclusion/exclusion criteria except for two occasions, when they entered patients into the study with HRSD-17 total scores of 13 at one of the diagnostic evaluations. While these events represented protocol violations, data from these two patients were left in the dataset since they did score ≥ 14 on the HRSD-17 at one point during the diagnostic phase, received A-CT, and in one case, was randomized to additional treatment in the continuation-phase of the study. Deleting this data from the dataset was considered wasteful and might have jeopardized the study's internal validity.

Acute-Phase Cognitive Therapy (Visit 401-420)

In this study, patients received 16-20 individual sessions of A-CT over a 12 to 14 week protocol, following Beck et al.'s (1979) treatment manual. As mentioned before, patients received A-CT twice weekly for four weeks (visit 401-408), at which time they were categorized as early or late responders. Early responders received sessions once a week for the final eight weeks (visit 409-416), while late responders continued receiving sessions twice weekly (visit 409-416) until the final four weeks of the acute-phase, when they received one session a week (visit 417-420). Overall, the 470 patients included in

this study received an average of 15.52 (*SD* = 5.80) sessions of A-CT, while early (*n* = 178) and late (*n* = 292) responders received an average of 15.74 (*SD* = 1.23) and 15.39 (*SD* = 7.30) sessions, respectively.

Treatment completion was operationalized as receiving ≥ 17 (if late responder) or 15 (if early responder) sessions of A-CT. Therefore, 79% of patients (*n* = 373) who met eligibility requirements and consented to begin A-CT completed the acute-phase. Ninety-seven patients dropped or withdrew from A-CT before receiving at least 17 (for late responders) or 15 (for early responders) sessions (see Table 5). Treatment non-completers received an average of 5.59 (*SD* = 4.86) sessions of A-CT.

Table 5

Reasons Patients did not Complete Acute-Phase Cognitive Therapy

Non-Completers of Acute-Phase Cognitive Therapy			
	UT South-western	WPIC	Total
Refused further contact	13	30	43
Sought alternative treatment	4	4	8
Hospitalized	0	1	1
Suicidal Intent	2	1	3
Non-compliant with protocol	5	10	15
Moved out of area	14	1	15
Became ineligible after enrolled	1	0	1
Withdrew by staff	5	6	11
Total	44	53	97

Note. UT Southwestern = The University of Texas Southwestern Medical Center at Dallas; WPIC = Western Psychiatric Institute and Clinic at Pittsburgh.

Three issues were raised when implementing the above treatment protocol. First, evaluators used the LIFE Psychiatric Treatment History (see p. 84 for description) to record when non-protocol psychosocial or pharmacological treatments were received during the acute-phase. Based on this information, a team of evaluators and the study psychiatrist categorized patients as receiving non-protocol treatment if their additional treatment: a) focused on reducing depressive symptom severity, b) targeted the patient, and c) was delivered at a therapeutic dose. After disagreements in categorizations were

resolved through group consensus, it was decided that 20 patients received additional, non-protocol treatment(s) during the acute-phase (see Table 6).

Table 6

Dose, Duration, and Purpose of Additional Treatment(s) Received by Study Patients during Acute-Phase Cognitive Therapy

Symptom	Additional TX	Dose [a]	Duration [b]	# of Patients [c]
Psychopharmacological Interventions				
Anxiety	Clonazepam	1 mg	15 wks	1
	Lorazepam	1 mg	1 wk	1
	Xanax	0.25 mg	1 wk	1
Depression	Prozac	40 mg	2 wks	1
Insomnia	Ambien	10 mg	1 wk	1
	Melatonin	1-5 mg	1-17 wks	2
	Seroquel	25 mg	2 wks	1
	Temazepam	30 mg	1 wk	1
	Trazodone	50 mg	7 wks	1
				Total = 10
Psychosocial Interventions				
Depression and anxiety	Online support group	23 sessions	4 wks	1
	Support group	17 sessions	17 wks	1
	Self-help therapy	18 sessions	18 wks	1
				Total = 3
Mood Altering Medications				
Pain relief and insomnia	Excedrin PM	77 mg	12 wks	1
	Tylenol PM	75-500 mg	4-9 wks	3
				Total = 4
Mood Altering Herbs				
Insomnia	Hapizen	845 mg	4 wks	1
	Unison	25 mg	2-16 wks	3
	Valerian Root	150-300mg	2 wks	1
Stress	Kava Kava	150-300 mg	4 wks	1
				Total = 6

[a] Average daily dosage in milligrams or number of session. [b] Duration was broken down into weekly increments. [c] Total number of patients who received additional treatment is > 20 because one patient received two additional treatments and one patient received three additional treatments.

Second, three patients received more sessions of A-CT then they should have. Specifically, one patient was incorrectly categorized as a late responder and received 20 sessions, when he or she should only have received 16. Two others were correctly categorized as early responders and still received 17 and 20 sessions, respectively. Data from patients who either received too many sessions of A-CT (n = 3) or additional, non-protocol treatment (n = 20) were left in the dataset. This was done to save data and describe staff behavior when implementing the study protocol.

Third, seven patients reported serious adverse events (see Table 7) that were not related to study treatment.

Table 7

Serious Adverse Events Reported to Data Safety and Monitoring Board during Acute-Phase Cognitive Therapy

Event Date	Report Date	Serious Adverse Event Reported
		WPIC (*n* = 3)
10/3/2000 – 10/5/2000	10/3/2000	Hospitalization due to worsening of depressive symptoms
11/2/2000	12/18/2000	Hospitalization due to patient being involved in automobile accident
7/27/2001	8/14/2001	Hospitalization due to decompensation while leaving abusive husband
		UT Southwestern (*n* = 4)
11/20/2001	11/26/2001	Hospitalization due to broken arm
7/11/2002	7/16/2002	Hospitalized to remove bone spur from neck
11/4/2002	10/28/2002	Hospitalized to remove benign cyst from wrist
7/25/2004 – 7/27/2004	8/2/2004	Hospitalization due to gall bladder attack

Note. UT Southwestern = The University of Texas Southwestern Medical Center at Dallas; WPIC = Western Psychiatric Institute and Clinic at Pittsburgh.

Psychoeducational Visits (Visit 701)

Psychoeducational visits occurred during the week before beginning A-CT and within seven days of completing session 11 (i.e. week 7-8 for early responders; week 6-7 for late responders). At these 30 minute visits, evaluators: a) asked questions about experiences in the study, b) informed patients about their risk of relapse/recurrence, c) reviewed study protocol, d) provided rationale for continuation treatment to increase

acceptance of randomization and continued participation when less symptomatic, and e) collected self-report questionnaires.

Blind Evaluation (Visit 101)

Blind evaluations occurred within seven days of the end of the acute phase, every four months for the subsequent 32 months, when a patient dropped or withdrew from the study, and any time a relapse or recurrence occurred. This study only used data from the post-A-CT blind evaluation (visit 101) because of its focus on the acute-phase. The primary purpose of this two to three hour visit was to assess the patient's response to A-CT by: a) determining whether DSM-IV MDD criteria were met, b) evaluating symptom severity, and c) examining changes in symptoms/syndromes since the follow-up diagnostic evaluation. To accomplish these goals, independent evaluators blinded to continuation-phase cell assignment assessed patients with the HRSD-17 and SCID-I.

Evaluators also completed the three components of the Longitudinal Interval Follow-up Evaluation (LIFE-II; Keller et al., 1987), including the Psychiatric Status Rating (PSR), Psychiatric Treatment History, and Psychosocial Interview (see LIFE section on p. 82 for details). This measure allowed evaluators to retrospectively assess patients' psychiatric symptoms, treatment history, and psychosocial functioning during the month of diagnostic evaluations and three months of A-CT. Again, slight differences existed between study sites regarding what evaluator completed the LIFE. At WPIC, evaluators blinded to continuation-phase cell assignment completed all three parts of the LIFE. At UT Southwestern, blinded evaluators only completed the LIFE PSR, while evaluators with knowledge about whether or not patients were randomized to

continuation CT or continuation medication (they were still blinded to type of medication received) completed the LIFE Psychiatric Treatment History and Psychosocial Interview.

The post-A-CT blind evaluation also served to continue informed consent and provide clinical care to patients who did not respond to A-CT, dropped out, or withdrew. If patients were non-responders to A-CT (see p. 95 for definition), evaluators discussed treatment options and facilitated referral to the most appropriate treatment. If patients dropped out or withdrew from A-CT and did not return to the clinic for their post-A-CT blind evaluation, evaluators made treatment referrals over the phone and by mail. To promote data collection at blind evaluations, evaluators made extensive efforts to reach patients. For example, evaluators offered to do evaluations over the telephone if patients moved, and patients were compensated $20 for each blind evaluation.

Measures

Patients completed a battery of instruments before, during, and after A-CT (see Table 8). A detailed description of each measure and their psychometric properties was provided below.

Table 8

Schedule of Data Collection

Weeks of A-CT	-2	-1	1	2	3	4	5	6		7	8	9	10	11	12	13
Phase of Study	Diagnostic Evaluation	Diagnostic Follow-up	A-CT						Psycho-educational Session	A-CT						Blind Evaluation
Visit Number	002	003	401-410 (early responder) 401-412 (late responder)						701	411-416 (early responder) 413-420 (late responder)						101
			Psychiatric Diagnosis & Treatment History													
Demographic Questionnaires	x															
SCID-I	x															x
SCID-I: MDD	x					x					x				x	x
LIFE																x
			Depressive Symptom Severity													
HRSD-17	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
BDI	x		x	x	x	x	x	x	x	x	x	x	x	x	x	x
IDS-SR	x		x	x	x	x	x	x	x	x	x	x	x	x	x	x
			Psychosocial Functioning													
DYS [a]		x	x						x							x
GAF	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
RIFT																x
SAS-SR	x		x						x							x
Q-LES-Q [a]		x														x

Note. A-CT = acute-phase Cognitive Therapy; BDI = Beck Depression Inventory; DYS = Dyadic Adjustment Scale; GAF = Global Assessment of Functioning Scale; HRSD = Hamilton Rating Scale for Depression-17; IDS-SR = Inventory for Depressive Symptomatology – Self-Report; LIFE = Longitudinal Interval Follow-up Evaluation; RIFT = Longitudinal Interval Follow-up Evaluation – Range of Impaired Functioning Tool; Q-LES-Q = Quality of Life Enjoyment and Satisfaction Questionnaire; SAS-SR = Social Adjustment Scale-Self Report; SCID-I = Structured Clinical Interview for DSM-IV; SCID-I: MDD = Structured Clinical Interview for DSM-IV: Major Depressive Disorder Section.

[a] At WPIC, the DYS and Q-LES-Q were completed at the diagnostic evaluation (visit 002)

Measures of Psychiatric Diagnosis and Treatment History

Diagnostic Questionnaires: Patients completed questionnaires that provided information regarding their age, race, marital status, education, employment, and family history of psychiatric disorders. Evaluators also completed questionnaires to clarify the nature of each patient's MDD, including length of the current episode, age of onset, number of episodes, and subtyping.

Structured Clinical Interview for DSM-IV Research Version, Patient Edition with Psychotic Screen (SCID-I): The SCID-I (First et al., 1996) was developed as a structured interview to assess the degree to which patients met criteria for DSM-IV diagnoses (American Psychiatric Association, 1994). The SCID-I has shown acceptable inter-rater reliability, both overall (kappa = 0.74) and when only assessing criteria for MDD (kappa = 0.72; Riskind, Beck, Berchick, Brown, & Steer, 1987). While the complete SCID-I was only administered at the initial diagnostic evaluation, the current MDE section was administered again at week four, eight, and 12 of A-CT and at the post-acute-phase blind evaluation (visit 101) to track patients' diagnostic status for MDD. In this study, inter-rater agreement equaled 0.87 across raters at both sites on the current MDE section of the SCID-I.

Longitudinal Interval Follow-up Evaluation (LIFE): The LIFE (Keller et al., 1987) was developed as a semi-structured interview to assess the longitudinal course of psychiatric disorders. It was comprised of three sections (i.e., Psychiatric Status Ratings, Psychosocial Interview, and Treatment History) that supplemented cross-sectional batteries and improved understanding of what happened between assessments in long-term studies. Administered in this study every four months, experienced raters asked

patients to recall their symptoms, treatment history, and psychosocial functioning retrospectively over weekly and monthly periods. Regarding inter-rater reliability, Keller et al. reported a median coefficient of 0.77 for the psychosocial functioning items and a range of 0.70 to 0.90 for the psychiatric status ratings. Also, the LIFE demonstrated good test-retest reliability and external validity (Warshaw, Keller, & Strout, 1994), and training programs have successfully prevented "rater-drift" (Warshaw, Dyck, Allsworth, Stout, & Keller, 2001).

With the *LIFE Psychiatric Status Ratings* (PSR), raters tracked patient's symptoms over the past four months using weekly intervals. For MDD and other "major affective disorders", a six-point scale was used to record duration and severity of symptoms, where scores of 5-6 indicated the presence of a Major Depressive Episode; 4-2 indicated residual symptoms; and one indicated a return to normal self (Keller et al., 1987). Raters used a 3-point scale to assess chronic depression and other psychiatric disorders, where three indicated the presence of a disorder; two indicated residual symptoms; and one indicated a return to normal self.

On the *LIFE Psychosocial Interview*, raters recorded patient's psychosocial functioning retrospectively over the past four months along monthly intervals. Raters asked patients to describe their work, family relationships, friendships, sexual functioning, recreation, satisfaction with overall functioning, and global social adjustment during the "worst" week in each monthly interval. As a result, scores reflected the week during which the most impairment in psychosocial functioning took place in each respective month. Queries about work allowed patients to identify their primary roles and rate their functioning as an employee, homemaker, and student. All questions were

rated on a 5-point scale. In this study, select items from the LIFE Psychosocial Interview were summed and scored using the RIFT protocol (see RIFT section below for details, p. 86).

To complete the *LIFE Psychiatric Treatment History*, raters recorded information regarding patients' treatment history over the last four months in weekly intervals, including any psychopharmacological and psychosocial treatments. If psychopharmacological treatment was received, raters coded the type and dose of medication, dates it was taken, and the diagnosis or symptom for which treatment was sought. For psychosocial treatments, raters recorded the type of treatment sought, the number of weekly sessions, and the psychological diagnosis for which treatment was sought. This information was used to determine which patients received additional, non-protocol treatment during A-CT (see Table 6).

Measures of Psychosocial Functioning

Dyadic Adjustment Scale (DYS): The 32-item, self-report DYS (Spanier, 1976) measured global relationship functioning and satisfaction. It was administered at the diagnostic evaluation (visit 002) at WPIC, the diagnostic follow-up evaluation (visit 003) at UT Southwestern, week one (visit 401) and seven (visit 701) of A-CT at both sites, and the post-A-CT blind evaluation (visit 101) at both sites. The score obtained during the diagnostic phase served as the baseline. Spanier (1976) reported scores on the scale ranged from 0 to 151, with mean scores of 114.8 (*SD* = 17.8) for married couples and 70.7 (*SD* = 23.8) for divorced couples. Given these findings, some have used a total score of 100 as the cutoff between populations with or without marital discord (i.e.,

Beach & O'Leary, 1992). The DYS was used in this study to examine the convergent and discriminant validity of the RIFT, SAS-SR, and index of psychosocial functioning.

Spanier (1976) found the DYS demonstrated high internal consistency reliability (alpha = 0.96) and convergent validity with the Locke-Wallace Marital Adjustment Test (r = 0.86; MAT; Locke & Wallace, 1959), a pre-existing measure that was adopted into the DYS. In this study, the median internal consistency reliability for the DYS equaled 0.95 (range equaled 0.94 to 0.95). Also, this study found the DYS showed acceptable levels of convergent validity with the SAS-SR (r = -0.39) but not with the GAF (r = 0.11) or RIFT (r = -0.14) (see Appendix G for details of these analyses).

Global Assessment of Functioning (GAF): The clinician-report GAF (American Psychiatric Association, 1994) measured overall functioning across a range of one to 100, with descriptive anchors at 10 point increments. To score the GAF, clinicians took into account both the patient's symptom severity and overall level of occupational, social, and psychological functioning. According to the DSM-IV-TR, a score ≤ 70 was needed to indicate some impairment in psychosocial functioning (American Psychiatric Association, 2000). In this study, the GAF was administered at both diagnostic evaluations, every week of A-CT, and at the post-A-CT blind evaluation, with the score obtained at the diagnostic evaluation (visit 002) serving as the baseline. As with the DYS, the GAF was used in this study to examine the convergent and discriminant validity of the RIFT, SAS-SR, and index of psychosocial functioning.

The GAF has demonstrated inter-rater reliability coefficients ranging from 0.80 (Goldman et al., 1992) to 0.92 (Hilsenroth et al., 2000) and a test-retest reliability coefficient of .69 (Spitzer & Forman, 1979). However, results regarding the GAF's

convergent validity have been mixed. Some show the GAF demonstrated acceptable convergent validity with other measures of psychosocial functioning (Skodol, Link, Shrout, & Horwath, 1988), while others suggested the GAF was better related to scales of psychiatric symptoms rather than psychosocial functioning (Hilsenroth et al., 2000). This study did not resolve this issue, as the GAF significantly correlated with both measures of psychosocial functioning (i.e., r = -0.26 with RIFT and -0.38 with SAS-SR) and depressive symptom severity (i.e., r = -0.36 with BDI, -0.45 with HRSD-17, and -0.35 with IDS-SR) (see Appendix G for details of these analyses).

Longitudinal Interval Follow-up Evaluation – Range of Impaired Functioning Tool (RIFT): Leon et al. (1999) developed the RIFT (see Appendix F for RIFT items and scoring procedure) as a brief measure of psychosocial functioning to gather information from four areas: interpersonal functioning, occupation functioning, recreational functioning, and satisfaction with overall functioning. The four-item RIFT produced a total score ranging from four to 20, with higher scores equaling higher impairment in psychosocial functioning. Since Leon et al. developed the RIFT with items from the LIFE Psychosocial Interview, the RIFT could be used to create total scores from data obtained using the LIFE-Psychosocial Interview. In this study, clinicians used the LIFE Psychosocial Interview to rate patients' usual level of functioning over the four months preceding the post-A-CT blind evaluation (visit 101). From this data, four RIFT total scores were extracted that coincided with the month preceding treatment and each month of A-CT. The score obtained for the month of diagnostic evaluations served as the baseline.

From a psychometric standpoint, Leon et al. (1999) found the RIFT demonstrated acceptable construct and concurrent validity (with the GAF), as well as internal consistency (median = 0.82) and inter-rater reliability (ICC = 0.94). Regarding predictive validity, Leon et al. also found the RIFT significantly distinguished populations with or without depression, with total scores averaging 14 and nine, respectively. This study found the RIFT demonstrated acceptable internal consistency reliability (median alpha = 0.68; range = 0.60 to 0.75), convergent validity with the SAS-SR and GAF, and discriminant validity with the BDI, HRSD-17, and IDS-SR (see Appendix G for details of these analyses).

A weakness of the RIFT was that it relied heavily on patients' ability to accurately recall events from their past. However, it was important to note that the other measures of psychosocial functioning used in this study (e.g., DYS, GAF, SAS-SR, and Q-LES-Q) also depended on patients' memory, just for substantially shorter time frames. The RIFT set itself apart from these measures by assessing psychosocial functioning longitudinally, whereas the other measures of psychosocial functioning in this study were cross-sectional and referenced the time period in their instructions (e.g., SAS-SR and Q-LES-Q asked about the past two weeks, the DYS and GAF asked about the past week).

Quality of Life Enjoyment and Satisfaction Questionnaire (Q-LES-Q): The 93-item, self-report Q-LES-Q measured "the degree of enjoyment and satisfaction experienced by patients in various areas of daily functioning" (p. 321; Endicott et al., 1993). These areas of functioning included physical health, subjective feelings, leisure activities, social relationships, general activities, work, household duties, and school/course work. Patients rated their enjoyment or satisfaction over the past week

along 5-point scales, with higher scores showing better functioning. Items were summarized into subscales, which were converted into percentages by dividing the raw score by the maximum possible subscale score. The Q-LES-Q was administered at the diagnostic evaluation (visit 002) at WPIC, the diagnostic follow-up evaluation (visit 003) at UT Southwestern, and the post-A-CT blind evaluation (visit 101) at both sites. The score obtained during the diagnostic evaluation served as the baseline.

While the Q-LES-Q contains some content unrelated to psychosocial functioning (e.g., physical health and subjective feelings), it is often used to measure the construct (Hirschfeld et al., 2000). Consequently, it was used in this study to examine the convergent and discriminant validity of the RIFT, SAS-SR, and index of psychosocial functioning. Endicott et al. (1993) found the Q-LES-Q demonstrated good test-retest (*r* ranged from 0.63 to 0.89) and internal consistency reliability (alpha ranged from 0.90 to 0.96). In this study, the total score of the Q-LES-Q demonstrated acceptable internal consistency reliability (median alpha = 0.91, range = 0.90 to 0.92), while subscales measuring psychosocial functioning showed convergent validity with similar subscales from the RIFT and SAS-SR (see Appendix G for details of these analyses).

Social Adjustment Scale-Self Report (SAS-SR): The 56-item, self-report SAS-SR measured "instrumental or expressive role performance in six major areas of functioning: work as a worker, housewife, or student; social and leisure activities; relationships with extended family; marital roles as a spouse, a parent, and a member of the family unit" (p. 1112; Weissman & Bothwell, 1976). Patients rated their functioning along 5-point scales, with higher scores indicating more impairment. These scores were averaged to

obtain subscale and total scores. Of note, the total score only took into account the subscale score from the work items pertaining to the primary work role.

In this study, the SAS-SR was administered at the diagnostic evaluation (visit 002), first (visit 401) and seventh (visit 701) week of A-CT, and post-A-CT blind evaluation (visit 101). The score obtained at the diagnostic evaluation served as the baseline. In 2001, Weissman et al. found total scores for individual diagnosed with MDD averaged 2.5, compared to an average total score of 1.7 for individual with no psychiatric disorder.

The SAS-SR has demonstrated acceptable levels of internal consistency (coefficient = 0.74) and test-retest reliability (coefficient = 0.80; Weissman, Prosoff, Thompson, Harding, & Myers, 1978). Weissman and colleagues (1978) also found the SAS-SR demonstrated acceptable predictive validity by differentiating psychiatric patients and community controls, as well as acute and recovered depressive patients (Weissman & Bothwell, 1976). Moreover, the SAS-SR has high convergent validity with the SF-36 Social Functioning Subscale (r = 0.42) and Social Adaptation Self-Evaluation Scale (r = 0.57; Weissman et al., 2001). In this study, the SAS-SR demonstrated acceptable internal consistency reliability (median alpha = 0.77, range = 0.76 to 0.78), convergent validity with the DYS, GAF, and RIFT, and discriminant validity with the HRSD-17 (see Appendix G for details of these analyses).

Measures of Depressive Symptoms

Beck Depression Inventory (BDI): The 21-item, self-report BDI measured the behavioral manifestation of depression over the past week (Beck et al., 1961). In this

study, patients completed the BDI at the diagnostic evaluation (visit 002), every week of A-CT, and at the post-A-CT blind evaluation. The score obtained at the diagnostic evaluation served as the baseline. Patients rated their depressive symptoms severity on a 4-point scale, ranging from 0 to 3. With the total score ranging from 0 to 63, depression could be categorized as minimal (0 to 10), mild to moderate (10 to 18), moderate to severe (19 to 29), or severe (> 29). The BDI-II (e.g., Beck, Steer, & Brown, 1996) was not used in this study due to its cost and its high correlation with the BDI (r = 0.93; Dozois, Dobson, & Ahnberg, 1998).

The BDI has demonstrated acceptable levels of convergent validity (correlations range from 0.61 to 0.87 with the HRSD-17), internal consistency reliability (*M* = 0.87), and test-retest reliability (*M* = 0.60) (Beck, Steer, & Garbin, 1988). In this study, it again showed acceptable internal consistency reliability (median alpha = 0.88, range = 0.83 to 0.92).

Hamilton Rating Scale for Depression (HRSD-17): With the HRSD-17, clinicians assessed depressive symptom severity over the past week along three (0 to 2) or five-point scales (0 to 4) (Hamilton, 1960). These scores were summed to produce a total score indicative of severe (> 24), mild (< 17), or no depression (< 6). In this study, scores were used from both diagnostic evaluations (visit 002 and 003), week one and seven of A-CT, and the post-A-CT blind evaluation. While total scores obtained at the diagnostic evaluation (visit 002) served as treatment baseline when using the index of depressive symptom severity, total scores from the diagnostic follow up (visit 003) where used when determining treatment response. Experienced evaluators completed the

HRSD-17 at the diagnostic evaluation, diagnostic follow-up, and the post-A-CT blind evaluation, while patient therapists completed the HRSD-17 during A-CT.

The HRSD-17 has demonstrated acceptable inter-rater reliability (r = 0.85; Clark & Watson, 1991), internal consistency reliability (alpha = 0.89; Rush, Gullion, Basco, Jarrett, & Trivedi, 1996), and convergent validity with the BDI (r = 0.75; Faravelli, Albanesi, & Poli, 1986). In this study, it again demonstrated acceptable internal consistency reliability (median alpha = 0.74; range = 0.63 to 0.84), except at the diagnostic evaluation (alpha = 0.51). HRSD-17 data from the diagnostic evaluation were kept, however, due to the restricted variability in the dataset, which could lower alpha estimates, and its high correlation with the BDI (r = 0.43) and IDS-SR (r = 0.53) at that time point.

Inventory for Depressive Symptomatology – Self-Report (IDS-SR-30): The 28-item "IDS-SR was designed to measure specific signs and symptoms of depression in both inpatients and outpatients", including those that were vegetative, cognitive, emotional, and endogenous (p. 67; Rush et al., 1986). Both self-report and clinician rated version of the IDS existed. Patients self-reported depressive symptom severity over the past week along four-point scales (0 to 3), with higher scores indicating greater impairment. Items were summed to produce a total score ranging from 0 to 78. In this study, patients completed the IDS-SR at the diagnostic evaluation (visit 002), every week of A-CT, and at the post-A-CT blind evaluation. The score obtained at the diagnostic evaluation served as the baseline.

The IDS-SR has shown adequate internal consistency reliability (alpha = .92; Trivedi et al., 2004) and concurrent validity with the BDI (r = 0.78) and the HRSD-17 (r

= 0.67; Rush et al., 1986). Rush et al. also found the IDS-SR demonstrated predictive validity, as total scores significantly differentiated individuals with depression (*M* = 36.5), minor depression (*M* = 21.8), or no depression (*M* = 2.1). In this study, the IDS-SR demonstrated acceptable internal consistency reliability (median alpha = 0.87, range = 0.82 to 0.92) and convergent validity with the BDI (*r* = 0.80).

Operationalization of Study Variables

Psychosocial Functioning

Previous research has shown that instruments measure psychosocial functioning differently, such that the type of instrument largely determines how much change occurs in psychosocial functioning during acute-phase treatment (Dunn et al., 2007). Therefore, an index of psychosocial functioning was created to: a) avoid spurious findings attributable to differences across instruments rather than treatment effects, b) ensure the assessment of psychosocial functioning was comprehensive in terms of content relevant to each domain, and c) better control for shared method variance. This index consisted of total scores from the RIFT and SAS-SR, combining data from multiple methods (i.e., self-report and clinician-rated) and time frames (i.e., cross-sectional and longitudinal). Specifically, while the SAS-SR was collected cross-sectionally and recorded psychosocial functioning data over two week intervals, the RIFT was collected longitudinally, allowing clinicians to rate patient psychosocial functioning over monthly intervals. As a result, scores on the index of psychosocial functioning represented the month of diagnostic evaluation and the first, second, and third month of A-CT. The

index did not include data from the: a) DYS because over half of the sample was single (i.e., 57.9%) and would not complete this instrument, b) Q-LES-Q because it included information that did not relate to psychosocial functioning (i.e., physical functioning and subjective feelings), and c) GAF because it also accounted for symptom severity.

Before creating the index of psychosocial functioning, the psychometric quality of the RIFT and SAS-SR was first examined. This initial step was especially important to justify the inclusion of RIFT data on the index, since researchers have recommended that retrospective measures not be used in structural equation modeling because of potential memory biases (Cole & Maxwell, 2003). Specifically, the RIFT and SAS-SR were required to demonstrate: a) internal consistency reliability, b) construct validity with baseline measures of psychosocial functioning, and c) sufficient discriminate validity with measures of depressive symptom severity. Since the RIFT has shown acceptable psychometric properties previously (e.g., Eisen et al., 2006; Judd et al., 2005; Solomon et al., 2004; Vittengl et al., 2008), it was expected to meet these requirements.

As a second step, total scores from the RIFT and SAS-SR were converted into a common metric (i.e., *t*-scores), using the following formula (e.g., Minium, King, & Bear, 1993). Total scores at each assessment point were standardized using their respective baseline mean and standard deviation.

$$X_T = \frac{X_{\text{raw}} - \overline{X}_{\text{pre-treatment}}}{SD_{\text{pre-treatment}}} \times 10 + 50$$

Once total scores were standardized, they were averaged and again standardized to maintain the mean = 50 and standard deviation = 10. After this transformation, *higher scores on the index of psychosocial functioning indicated greater impairments*.

Depressive Symptoms

Since scales of depressive symptom severity varied in their degree of overlap with psychosocial functioning (e.g., Vittengl et al., 2004), an index of depressive symptom severity was also created using multiple measures and informant types. Specifically, total scores from the BDI, HRSD-17, and IDS-SR-30 were converted into *t*-scores, using the formula presented above. Total scores from each measure were then averaged to create a single index of depressive symptom severity and again standardized to maintain the mean = 50 and standard deviation = 10. For this index, *higher scores equated to greater depressive symptom severity*.

Of note, only total scores from the diagnostic evaluation, week one and seven of A-CT, and the post-A-CT blind evaluation were used in order to match available data from measures of psychosocial functioning. Also, since each measure in the index collected data over weekly intervals, scores on the index represented depressive symptom severity at the week of diagnostic evaluation, the first and seventh week of A-CT, and the week of the post-A-CT blind evaluation. Past research has validated the inclusion of BDI, HRSD-17, and IDS-SR data in an index of depressive symptom severity, showing these measures have convergent validity (Vittengl, Clark, Kraft, & Jarrett, 2005) and possess higher internal consistency reliability together (i.e., alpha coefficient = 0.95) than separately (e.g., Vittengl et al., 2004).

Early vs. Late Response

Early response was defined as a ≥ 40% reduction in HRSD-17 scores by the eighth session of A-CT compared to the HRSD-17 obtained at the follow-up diagnostic evaluation (visit 003). Conversely, non-early, or late, response was defined as a < 40% reduction in HRSD-17 scores by the eighth session of A-CT compared to the same baseline.

Response

Treatment response was defined at the end of A-CT (i.e., visit 101) as both the absence of MDD, as determined by the SCID-I Current MDE, and an HRSD-17 of ≤ 12. If visit 101 (i.e., post-A-CT blind evaluation) data was not available, a proxy definition of treatment response was used. This proxy definition used data from the last available HRSD-17 to determine symptom severity and last two available IDS-SRs to determine the presence or absence of MDD. With the proxy definition, both treatment completers and non-completers were categorized as responders or non-responders to A-CT.

Study Design: Prerequisites to Analysis

Examining the Effects of "Extraneous" Variables

Before carrying out analyses with this data set, two issues warranted attention. First, the study design contained two embedded factors: study site and early vs. late

response to A-CT. While conducting the study at two sites could increase heterogeneity, and therefore external validity of the sample, differences across sites could also influence study results. Similarly, the categorization of patients as early vs. late responders introduced another factor into the study design that could affect results. If baseline differences existed across sites or response types, significant findings could reflect site characteristics or unique attributes of early vs. late treatment responders rather than treatment effects. Consequently, if significant differences were found in baseline levels of depressive symptom severity or psychosocial functioning across these two factors, steps were taken to better isolate treatment effects. For instance, these potential sources of confounding variability were controlled by entering them as factors in the analysis of variance (ANOVA) models, thereby isolating their main and/or simple effects (e.g., interaction between study site and time in treatment).

Missing Data Confirmation

A systematic procedure was needed to confirm the study's missing data and account for it during statistical analyses. To ensure missing data were truly missing, all missing data points were double checked with visit logs that tracked every session of A-CT. Data were considered truly missing if the corresponding visit never occurred. If missing data were reported for sessions of A-CT that actually took place, the missing data log was searched to see if the data were reported as missing for other reasons (e.g., incomplete questionnaire, error in filling out the questionnaire, etc.). If the missing data were not reported in the missing data log, the raw data was searched. If it turned out that the data were available, they were entered into the statistical dataset, but if the data were

discovered to be missing, the missing data log was updated. As a result of this process, all missing data in the study's dataset were verified as truly missing.

To account for missing data in preliminary aims four and five and the primary aim, a likelihood-based, mixed-effects model was employed, with the underlying assumption that observed factors accounted for the missing data rather than unobserved factors. This model avoided the biases often associated with the last observation carried forward (LOCF) approach because it was carried out without any manipulation of the missing data points (Mallinckrodt, Watkins, Molenberghs, & Carroll, 2004). Consequently, missing data points in these analyses were not replaced with the last observation or an average of available data.

For preliminary analyses one through three, however, only available data were used. This was justified given the nature of analyses in preliminary aim one (i.e., psychometric properties rather than group comparisons), and the timing of analyses in preliminary aims two and three (i.e., treatment baseline, when all or nearly all data was available for each measure). Also, research suggested the manipulating of missing data points (e.g., replacing them with the last observation or average of available data) could bias results (Mallinckrodt et al., 2004).

Changes during this audit process (e.g., adding data that were previously thought missing) were only made to the statistical dataset, which represented a portion of the master dataset collected by the primary investigators for NIMH R01 grants MH 58397 and MH 69619 (Principal investigator: Robin B. Jarrett, Ph.D.) and MH 58396 and MH 69618 (Principal investigator: Michael E. Thase, M.D.). While the statistical dataset

consisted of acute-phase data for 470 patients who were no longer active in A-CT by August 10, 2007, patient recruitment for the master dataset ended on August, 1, 2008, with data collection ending approximately 36 months thereafter.

Data Quality and Monitoring

Clinicians and evaluators were trained and monitored to maintain a high level of diagnostic reliability across sites. To accomplish this, project coordinators at both sites completed inter-site reliability studies on the SCID-I Current MDE and the HRSD-17 (see Appendix H for inter-rater reliabilities across sites). Specifically, all clinicians and evaluators met quarterly to co-rate videotapes containing the HRSD-17 and SCID-I Current MDE interviews. Videotaped interviews were exchanged between sites, such that two tapes from each site were rated each year. All ratings were entered into a reliability database at each site and shared via a secure website. Ratings from senior evaluators functioned as the "gold standard" with which other ratings were compared. Any evaluator who had HRSD-17 ratings that differed from the gold standard by ≥ four, or who disagreed on DSM-IV diagnoses, completed additional ratings and training as needed.

All study data was either entered by hand or scanned. To ensure the quality of the data, research staff double-keyed hand-entered data and resolved discrepancies between the first and second entries by consensus, using the original data. Also, both sites routinely ran audit programs to identify missing data or coding inconsistencies (e.g., incorrectly entered data) that required data editing. After the principal investigators approved appropriate data modifications, the data edits were appended to the master

database at each site then uploaded to a protected website for data sharing and downloading for statistical analyses.

To ensure the safety of the data, the research team saved the database on a protected website with restricted access to select study personnel. The principal investigators were notified each time data was uploaded from the website. The research team also made backups of the database and stored it on disks in a fireproof safe, a hard drive that was not connected to the Internet, and the O-drive, which was backed up nightly. Revisions to this dissertation were also saved nightly to the O-drive, weekly to a flash drive, and monthly to a hard drive on a different campus.

Research Questions, Hypotheses, and Statistical Analyses

Preliminary Aim 1

Research Question: What is the psychometric quality of the indices of psychosocial functioning and depressive symptom severity?

Rationale: Before investigating how and when change occurs in psychosocial functioning in relation to depressive symptom severity, one first needed to show that these two constructs were measured reliably and validly. Past research supported the use of the index of depressive symptom severity, showing: a) included instruments (e.g., BID, HRSD-17, and IDS-SR) had internal consistency reliability and convergent validity (Vittengl et al., 2005); and b) the index itself demonstrated internal consistency reliability and sensitive to changes during A-CT, with pre- to post-A-CT changes on the index

equaling an effect size of 1.55 (Vittengl et al., 2004). Results from the current study sought to replicate these findings.

The index of psychosocial functioning, on the other hand, had not previously been used in the literature. For this reason and due to research showing measures of psychosocial functioning assessed the construct differently (e.g., Dunn et al., 2007; Weissman et al., 2001), this study needed to first justify the inclusion of RIFT and SAS-SR data on the index of psychosocial functioning. This sub-analysis showed the RIFT and SAS-SR had sufficient internal consistency reliability and convergent and discriminant validity to be included on the index. Second, the index of psychosocial functioning, itself, needed examination to determine its reliability, validity, and sensitivity to pre- to post-A-CT changes. To improve study continuity, sub-analyses done to evaluate the psychometric properties of the RIFT and SAS-SR were presented in Appendix G. Steps taken to evaluate the psychometric quality of the index of psychosocial functioning and replicate past findings with the index of depressive symptom severity appeared below.

Research Questions Operationalized:

1) Do the indices of psychosocial functioning and depressive symptom severity have acceptable internal consistency reliability at treatment baseline and the beginning, middle, and end of A-CT?
2) Does the index of psychosocial functioning show convergent validity by significantly correlating with other commonly used measures of psychosocial functioning (e.g., DYS, GAF, and Q-LES-Q) at treatment baseline and the beginning, middle, and end of A-CT?

3) Does the index of psychosocial functioning show discriminant validity by not being highly correlated (i.e., $r < .50$; Cohen, 1988) with the BDI, HRSD-17, IDS-SR, and index of depressive symptom severity at treatment baseline and the beginning, middle, and end of A-CT?

4) Do the indices of psychosocial functioning and depressive symptom severity show sensitivity to changes during A-CT by having large pre- to post-A-CT standardized mean gain effect sizes (i.e., $d \geq 0.80$; Cohen, 1988)?

Hypotheses:

1) – 4) Given past research (Vittengl et al., 2004), it was predicted that the index of depressive symptom severity would show acceptable internal consistency reliability and a large pre- to post-A-CT standardized mean gain effect size. However, no hypotheses were generated for analyses with the index of psychosocial functioning, as these analyses were exploratory.

Statistical Analyses: To test hypothesis one, Cronbach's Alpha was calculated for the indices of psychosocial functioning and depressive symptom severity across this study's four assessment points, which represented treatment baseline and the beginning, middle, and end of A-CT. For the index of psychosocial functioning, scores represented the month of diagnostic evaluations and the first, second, and third month of A-CT, and for the index of depressive symptom severity, scores represented the week of diagnostic evaluation, the first and seventh week of A-CT, and the week of the post-A-CT blind evaluation. To test hypotheses two and three, Pearson product-moment correlation coefficients correlated the index of psychosocial functioning with the index of depressive

symptom severity and total scores from the DYS, GAF, Q-LES-Q, BDI, HRSD-17, and IDS-SR across these four assessment points. Findings were presented in table format to facilitate interpretation.

For hypothesis four, standardized mean gain effect sizes were calculated by subtracting pre- and post-A-CT means on the indices of psychosocial functioning and depressive symptom severity and dividing these numbers by their respective pooled variances. Effect sizes were used to compare the magnitude of change in these two constructs, as captured by their respective indices, to changes previously documented in randomized controlled trials (RCTs) investigating A-CT. If effect sizes across studies were equivalent, it could be concluded that the changes observed in this study were comparable to exposure to A-CT, and not likely due to regression to the mean or some other confounding variable.

Preliminary Aim 2

Research Questions: Do the demographic and clinical characteristics of study patients differ across study sites?

Rationale: This open trial was conducted at two sites: UT Southwestern and WPIC. Differences could arise across sites that affect the interpretation of results. This preliminary aim determined the extent to which baseline differences existed between sites across demographic and clinical characteristics. If a site difference was substantiated, steps were taken to control for this source of variability to better isolate treatment effects in preliminary aims four and five.

Research Question Operationalized: Do the demographic and clinical characteristics of study patients significantly differ at baseline across study sites?

Hypothesis: No hypotheses were generated, as these were exploratory analyses.

Statistical Analyses: For categorical variables, data was presented in terms of percent frequency and comparisons were made with the Pearson Chi-Square or Fisher exact test. With continuous variables, data was presented in terms of means and standard deviations and comparisons were made with independent group T-tests. Levene's test for equality of variances was run to determine the extent to which variances differed across groups. If variances were found to significantly differ, as was the case for age of onset, length of current episode, and HRSD-17 at diagnostic follow-up, independent groups T-tests were performed that did not assume equal variances. All comparisons were 2-tailed with a conservative alpha level of 0.05.

Preliminary Aim 3

Research Question: To what extent do baseline demographic and clinical characteristics differ across early vs. late responders to A-CT?

Rationale: This preliminary aim investigated the extent to which early and late responders differed across baseline demographic and clinical characteristics. If baseline differences in psychosocial functioning and depressive symptom severity were detected across response patterns, steps were taken to control for the impact of an early or late response to A-CT in order to better isolate treatment effects in preliminary aims four and five.

Research Question Operationalized: Do early and late responders to A-CT significantly differ across demographic and clinical characteristics at baseline?

Hypotheses: No hypotheses were generated, as these were exploratory analyses.

Statistical Analyses: For categorical variables, data was presented data in terms of percent frequency and comparisons were made with the Pearson Chi-Square and Fisher exact test. With the continuous variables, data was presented in terms of means and standard deviations and comparisons were made with independent group T-tests. A single chi-square analysis was also conducted to test the hypothesis that the response patterns observed at UT Southwestern and WPIC were independent of site location. Levene's test for equality of variances was again run to determine the extent to which variances differed across groups. If variances were found to significantly differ, as was the case for length of current episode, independent groups T-tests were performed that did not assume equal variances. All comparisons were 2-tailed with a conservative alpha level of 0.05.

Preliminary Aim 4

Research Question: To what extent does depressive symptom severity improve after exposure to A-CT compared to pre-treatment baseline?

Rationale: The primary aim of this study investigates how psychosocial functioning changes during A-CT for MDD, both independently and in relation to changes in depressive symptoms. Before this research question can be pursued, one must first show that depressive symptom severity does indeed change during A-CT. To accomplish this, analyses: a) compared scores on the index of depressive symptom

severity at the post-A-CT blind evaluation to scores taken at previous assessment points and b) determined the number of patients who experienced enough symptom improvement to be considered treatment responders. Steps were also taken in this analysis to ensure pre- to post-A-CT changes in depressive symptom severity were not due to the presence or absence of additional treatment or differences in embedded factors in the study design (i.e., study site and early vs. late response pattern).

Research Question Operationalized:

1) After accounting for confounding variables, is the index of depressive symptoms obtained after A-CT significantly lower than pre-treatment baseline?
2) What percentage of patients who terminate A-CT meet criteria for treatment response?

Hypothesis:

1) After confounding variables are controlled, the index of depressive symptom severity will be significantly lower after A-CT compared to pretreatment baseline.
2) Based on past similar research (Jarrett et al., 2001), approximately two-thirds of patients are expected to meet criteria for treatment response (see p. 95 for details of definition).

Statistical Analyses: To test hypothesis one, two steps were taken. First, if significant baseline differences existed between study sites and early vs. late responders to A-CT on measures of depressive symptom severity, two two-way repeated measures ANOVAs (one for each factor) were run to determine the extent to which these

embedded factors influenced pre- to post-A-CT changes on the index of depressive symptom severity. During this step, a two-way repeated measures ANOVA was also run to examine the impact of additional treatment on pre- to post-A-CT changes in the index of depressive symptom severity.

For these three analyses, the dependent variable was the index of depressive symptom severity, and the within-groups independent variable was "time in A-CT", which represented the index of depressive symptom severity over repeated assessments during A-CT (i.e., diagnostic evaluation, week one and seven of A-CT, and post-A-CT blind evaluation). Separate analyses were run for each between-groups independent variable, which was "study site", "early vs. late response to A-CT", or "presence or absence of additional treatment". If a significant interaction between "time in A-CT" and one of the between-group factors was noted, it could be concluded that pre- to post-A-CT changes in the index of depressive symptom severity were influenced by: a) where that treatment occurred (i.e., UT Southwestern or WPIC), b) an early vs. late response pattern to A-CT, and/or c) the presence or absence of additional treatment.

For the second step, a repeated-measures ANOVA was run to determine the extent to which the index of depressive symptom severity lowered from pre- to post-A-CT, controlling for all between-groups factors that significantly influenced this change in step one. This step facilitated an understanding of the relative impact of each potential confounding variable on pre- to post-A-CT changes in the index of depressive symptom severity, since it was possible that the influence of one between-group factor was accounted for by another.

All repeated-measures ANOVAs were based on maximum likelihood estimation, with a 2-tailed, conservative alpha level of 0.05, since these were exploratory analyses. Also, post hoc analyses (e.g., Bonferroni) and effect sizes (R^2) facilitated interpretation of results. For instance, if R^2 for the interaction between "time in A-CT" and one of the between-group factors was small (i.e., $\leq$ 0.02; Cohen, 1988), the inclusion of the between-group factor in the final ANOVA model would be questioned because it accounted for little variability in changes on the index of depressive symptom severity during A-CT. Effect sizes were calculated by converting *F*-values to *t*-values ($F = t^2$) and using the below formula to calculate R^2 (Lipsey et al., 2001):

$$R^2 = \frac{t^2}{t^2 + df}$$

In the above equation, *df* equaled *N* minus the number of groups (e.g., *df* of denominator minus *df* of numerator).

To test hypothesis two, the percentage of responders was calculated and compared to response rates from RCTs. If response rates in this study were equivalent to those from RCTs investigating A-CT, the inference that changes observed in this study were comparable to exposure to A-CT would be substantiated.

Preliminary Aim 5

Research Question: To what extent does psychosocial functioning improve after exposure to A-CT compared to pre-treatment baseline?

Rationale: Similar to preliminary aim four, before the primary aim can be pursued or interpreted, it is important to first show that psychosocial functioning does indeed change from pre- to post-A-CT. To do this, the current aim compared scores on the index of psychosocial functioning obtained during the last month of A-CT to scores obtained during previous assessment points. Also as before, steps were taken in this analysis to ensure pre- to post-A-CT changes psychosocial functioning were not due to the presence or absence of additional treatment or differences in embedded factors in the study design (i.e., study site and early vs. late response pattern).

Research Question Operationalized: When confounding variables are controlled, are scores on the index of psychosocial functioning measured after A-CT significantly lower compared to pre-A-CT baseline?

Hypothesis: After controlling for confounding variables, the index of psychosocial functioning will significantly decrease from pre- to post-A-CT.

Statistical Analyses: The statistical analyses for this preliminary aim were similar to those in preliminary aim four, with the only exception that the dependent variable changed. As a first step, if significant baseline differences existed between study sites and early vs. late responders on measures of psychosocial functioning, two two-way, repeated measures ANOVAs were run to determine the extent to which these embedded factors influenced pre- to post-A-CT changes on the index of psychosocial functioning. During this step, a two-way, repeated measures ANOVA was also run to examine the impact of additional treatment on pre- to post-A-CT changes on the index of psychosocial functioning.

For theses analyses, the dependent variable was the index of psychosocial functioning, and the within-group independent variable was "time in A-CT," which represented changes in the index of psychosocial functioning over repeated measurements during A-CT (i.e., month of diagnostic evaluations and first, second, and third month of A-CT). Separate analyses were run for each between-groups independent variable, which were "study site", "early vs. late response to A-CT", or "presence or absence of additional treatment". If a significant interaction between "time in A-CT" and one of the between-groups factors was noted, it was concluded that pre- to post-A-CT changes on the index of psychosocial functioning were influenced by: a) where treatment occurred (i.e., UT Southwestern or WPIC), b) an early vs. late response pattern to A-CT, and/or c) the presence or absence of additional treatment.

For the second step, a repeated-measures ANOVA was run to determine the extent to which the index of psychosocial functioning reduced from pre- to post-A-CT, controlling for all between-group factors that significantly influenced this change in step one. This step improved understanding of the relative impact of each confounding variable on pre- to post-A-CT changes on the index of psychosocial functioning, since it was possible that the influence of one between-group factor was accounted for by another.

As before, all repeated-measures ANOVAs were based on maximum likelihood estimation, with a 2-tailed, conservative alpha level of 0.05, since these were exploratory analyses. Also, post hoc analyses (e.g., Bonferroni) and effect sizes (R^2) facilitated interpretation of results.

Primary Aim

Research Question: During A-CT for MDD, to what extent are changes in variance unique to measures of psychosocial functioning mediated by changes in variance unique to measures of depressive symptom severity?

Rationale: Psychosocial functioning is an important outcome for acute-phase interventions of MDD (e.g., Keller, 2003; Thase, 2003). Acute-phase CT significantly reduces impairment in psychosocial functioning, but it is still unclear when and how these changes occur (e.g., Vittengl et al., 2004). New knowledge will be created if researchers address when and how psychosocial functioning changes during A-CT, both independently and in relation to depressive symptom severity.

Research Questions Operationalized:

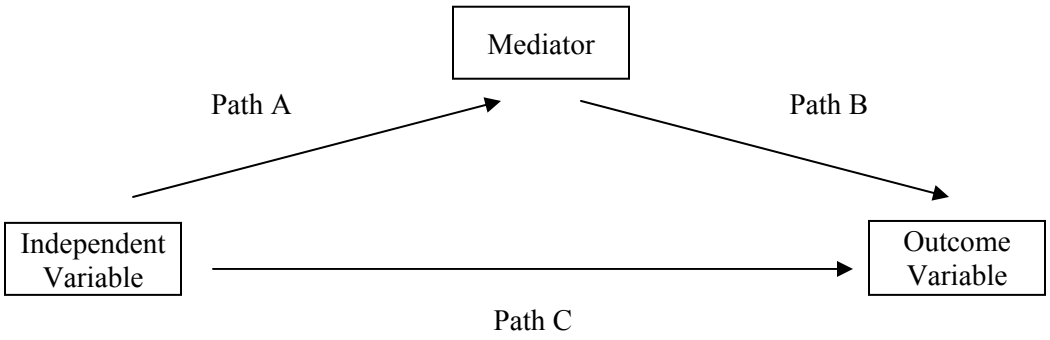
1) Does the structural equation model fit data accounting for:
    a. Covariation between psychosocial functioning and depressive symptom severity at treatment baseline and the beginning, middle, and end of A-CT,
    b. Shared sources of error variance across repeated measurements,
    c. Change in depressive symptom severity that occurs independently of change in psychosocial functioning,
    d. Change in psychosocial functioning that occurs independently of change in depressive symptom severity, and
    e. Potential mediating relationships between psychosocial functioning and depressive symptom severity?

2) Do changes in variance unique to measures of depressive symptom severity mediate changes in variance unique to measures of psychosocial functioning?

Hypotheses:

1) This structural equation model will show good fit with the data across four indices of fit.
2) Since past research shows change in depressive symptom severity precedes and accounts for change in psychosocial functioning across treatment settings (e.g., Judd, Akiskal, et al., 2000; Hirschfeld et al., 2002; Vittengl et al., 2004), it is expected that change in depressive symptom severity will partially mediate change in psychosocial functioning.

Statistics: In their seminal work, Baron and Kenny (1986) used the following diagram and claimed a variable could be considered a mediator if:



(a) variations in levels of the independent variable significantly account for variations in the presumed mediator (i.e., Path a),

(b) variations in the mediator significantly account for variations in the dependent variable (i.e., Path b), and

> (c) when Paths a and b are controlled, a previously significant relation between the independent and dependent variables is no longer significant, with the strongest demonstration of mediation occurring when Path c is zero (p. 1176; Baron & Kenny, 1986).

To test the mediating relationship in the primary aim, it was possible to run a series of multiple linear regressions based on the above criteria, where scores of depressive symptom severity and psychosocial functioning were in turn entered as predictors and criterion. These analyses would, in essence, provide information on how much variability was accounted for by the predictor (e.g., change in depressive symptom severity) in determining the criterion (e.g., change in psychosocial functioning). Since these analyses did not account for temporal precedence between variables, however, some have claimed that they would not provide sufficient evidence for or against a mediating relationship (Kraemer, Stice, Kazdin, Offord, & Kupfer, 2001). Consequently, if such an analysis plan was adopted, even if changes in one variable were shown to mediate changes in another, it remained a possibility that a mediational relationship also existed when variables were reversed.

In response to this limitation, Wilson et al. (2002) required a mediating variable to demonstrate:

> (a) that it measures some process occurring after the onset of treatment but before the outcome is determined,
>
> (b) that the process is correlated with treatment type, and
>
> (c) that some portion of the effect of treatment is explained by variation in that measure (p. 268).

In their study of the mechanisms of change in CT for Bulimia Nervosa, they followed three-steps to meet these requirements to establish a mediating relationship (Wilson et al.). First, they only investigated variables as potential mediators if they were measured after treatment onset and before some predetermined outcome, showing the variable changed during treatment but before the outcome. Second, they conducted a series of multiple linear regressions to establish a correlation between the potential mediators and treatment conditions. Finally, a second series of multiple linear regressions were run to see if the potential mediators predicted or influenced treatment outcome. If a variable met all of these conditions, it was shown to mediate treatment effects on a certain outcome. However, some have also found that this method was limited due to its use of linear regression, which as a statistic did not: a) allow variables to be entered as both predictor and criterion in the same analysis or b) control for inter-correlation between variables across data sets (Kline, 2005).

With the criteria for mediating variables in mind (e.g., Baron et al., 1986; Wilson et al., 2002) and in order to circumvent the above mentioned limitations, it was decided to test the primary aim of this dissertation with structural equation modeling. In what follows, efforts were made to identify the essential elements of structural equation modeling, describe how a model was justified, and specify how the structural equation model in Figure 4 (see p. 123) was set up to test the primary hypotheses.

Elements of Structural Equation Modeling with Path Analysis: The structural equation model designed to test the primary aim of this dissertation was called a path analysis (PA; Kline, 2005). In PA, squares or rectangles represented "manifest" variables that were directly measured, whereas circles or ellipses represented "latent" variables that

were not directly measured. In Figure 4, for example, each rectangle represented a variable that had been directly measured. By manipulating the relationships between manifest variables, researchers tested causal hypotheses based on their understanding of the literature.

Manifest variables in the model without hypothesized causes were called *exogenous*, while manifest variables with hypothesized causes were called *endogenous*. In Figure 4, the rectangle factors labeled "Depressive Symptoms: Baseline" and "Psychosocial Functioning: Baseline" were exogenous factors because their causes were not identified. All the other rectangle factors in Figure 4 were endogenous because they had identified causes represented by the arrows that point to them from other factors. These arrows signified causal pathways, or direct effects, that identified how much variability in the endogenous variable was explained by the causal variable(s), controlling for the effects of extraneous factors.

Circles with an uppercase D, known as *disturbances*, represented the effects of extraneous factors. Said differently, disturbances represented the variability in the endogenous variables that was not accounted for by the direct or indirect causal effects in the model. Circles, rather than squares, were used to graphically represent disturbances because they were not directly measured in the study. The arrows that point from the disturbances to the endogenous factors represented the direct effect of these unknown sources of variability on the factor. As seen in Figure 4, the effect of the disturbances on the endogenous variables were assigned a scale of 1.0, which specified the metric by which these unexplained sources of variance were presented. Variables that did not have identified causes (i.e., exogenous variables) were considered free to vary and covary, as

the study did not seek to identify their explained or unexplained sources of variance. A curved arrow that exited and reentered the exogenous variable represented this freedom from explained or unexplained sources of variability.

Evaluation of a Path Analysis Model: In order to be evaluated, the PA model in Figure 4 needed to meet three conditions. First, the number of parameters could not exceed the number of observations. Parameters were characteristics of the population that were estimated with sample statistics, or the total number of explained and unexplained sources of variation in the model. Specifically, the number of parameters in a PA model equaled the total number of observed or unobserved variances of exogenous variables, including disturbances, and direct effects on endogenous variables. For this study's model (see Figure 4), the number of parameters equaled 34, including eight variances for exogenous variables (two observed and six unobserved), 10 covariances between observed exogenous variables, and 16 direct effects.

The number of observations represented the number of parameters a model could estimate without exceeding its empirical bounds. In other words, if a model had more parameters than observations, it became mathematically impossible to "derive unique estimates for each parameter" (p. 101; Kline, 2005). The number of observations for a PA model was calculated with the following formula: $v(v + 1)/2$, where v equaled the number of manifest variables. As the proposed PA model in this study had eight manifest variables (i.e., rectangles), the number of observations equaled 36. The degrees of freedom for PA models (df_M) equaled the number of observations minus the number of parameters. Since the number of observations must exceed the number of parameters, if

the model was to be justified, the df_M had to be greater than or equal to zero. For the proposed PA model in Figure 4, $df_M = 36 - 34 = 2$.

Second, every exogenous variable that was not measured (e.g., disturbances) had to be given a scale so the computer program knew what metric to use in calculating its direct effect on the endogenous factors. As seen in Figure 4, the direct effect for each disturbance was assigned the scale of 1.0. Lastly, Kline (2005) recommended that PA models had sample sizes representing at least a 10:1 ratio to the number of model parameters. As mentioned before, the PA model in this study had 34 parameters, so the minimum sample size was 340. The sample size for this study was 470.

Since the PA model met the three criteria proposed by Kline (2005), it was evaluated using a two step process. First, model fit was determined by testing hierarchical models with four indices of fit that represented current best practice (McDonald & Ho, 2002). Model fit was defined as the degree to which the "predicted correlations and covariances equal their observed counterparts" (Kline, 2005). In other words, if the observed pattern of covariation in the model did not significantly differ from what was predicted through maximum likelihood estimation assuming multivariate normality, the model was said to fit the data. The four fit indices were the model chi-square (χ_m^2), root mean square error of approximation (RMSEA), comparative fit index (CFI), and normed fit index (NFI). With the model chi-square, a failure to reject the null hypothesis supported the PA model because larger values of chi-square indicated increasingly worse fit. For the RMSEA, values less than 0.10 represented good fit, and for the NFI and CFI, values at or above 0.90 indicated good fit. Multiple indices of fit were used because no one index could determine model fit alone (Kline, 2005).

Typically, when testing hierarchical models, researchers began with a just-identified model (i.e., number of parameters equals number of observations so that df_M = 0) and simplified it, trimming paths that did not significantly contribute to model fit (Kline, 2005). Evan though the model in Figure 4 was not just-identified (df_M = 2), this study took a similar approach. In other words, if the model in Figure 4 did not initially fit, paths were eliminated until fit indices fell within the above specified ranges, or a parsimonious model was created that fit the data well. If needed, paths were also built, or added to the model, to improve fit. The chi-square difference statistic (χ^2_D) was used to determine the statistical significance of each path deletion or addition. Essentially, the χ^2_D tested the null hypothesis that a revised model with one less, or more, path fit the data as well as the pre-existing model. A significantly positive χ^2_D indicated that the removal or addition of a path significantly improved model fit compared to the pre-existing model.

The second step involved evaluating parameter estimates of the model. To do this, direct effects were analyzed, together with their standard errors, to determine the degree to which endogenous variables were caused by the predicted pathways. Unstandardized direct effects were interpreted like path coefficients in multiple regression, meaning a one point increase in the predictor variable led to a quantified change in the endogenous variable. The significance of directs effects was calculated by dividing unstandardized parameter estimates by their respective standard error. This calculation produced a critical ratio that was then compared to the critical value of ± 3.2905, which represented an alpha of 0.001. The Bonferroni correction (i.e., α / number of comparisons; 0.05 / 16 direct effects, 8 variances, 10 covariances, and 6 indirect

effects; 0.05 / 40 = 0.001) was used to set alpha at 0.001 in order to reduce the increased probability of Type 1 error associated with multiple hypothesis testing.

Since LISREL 8.80 (Linear Structural Relationships) or AMOS 16.0 (Analysis of Moment Structures) did not report standard errors for indirect effects, the Sobel test was used to test the statistical significance of the indirect effects in the primary aim (Sobel, 1986). Specifically, the Sobel test calculated the standard error of an indirect effect through one mediator, with the following equation:

$$SE_{ab} = \sqrt{b^2 SE_a^2 + a^2 SE_b^2}$$

In this equation, a and SE_a represented the unstandardized parameter and standard error for path $X \rightarrow Y_1$, while and b and SE_b represented the unstandardized parameter and standard error for path $Y_1 \rightarrow Y_2$. The statistical significance of the unstandardized indirect effect (calculated as ab) was then tested by dividing it by the standard error (SE_{ab}) and comparing the critical ratio with the critical value of ± 3.2905, which corresponded to an alpha of 0.001. Complete mediating effects occurred when there were significant indirect effects without significant direct effects. For example, if the indirect effects that connected the variables "Psychosocial Functioning: Baseline", "Depressive Symptoms: Week 1", and "Psychosocial Functioning: Month 2" were both significant in Figure 4, while the direct effect between "Psychosocial Functioning: Baseline" and "Psychosocial Functioning: Month 2" was not, a complete mediating effect was supported. If this occurred, the variable "Depressive Symptoms: Week 1" mediated, or accounted for the relationship between the variables "Psychosocial Functioning:

Baseline" and "Psychosocial Functioning: Month 2". If the all three paths were significant, a partial mediating relationship was supported by the model.

All analyses described above were first run using LISREL 8.80 software and then replicated using AMOS 16.0 software.

How Path Analysis Tests the Mediational Relationship in Our Primary Aim: Frequently, constructs in the social sciences (e.g., intelligence, competitiveness, depression, etc.) are not directly observable. However, researchers attempt to measure these constructs indirectly and make inferences about them with varying degrees of success. In the cases of depressive symptom severity and psychosocial functioning, instruments in this study varied in terms of reliability and validity. Consequently, it was unreasonable to expect a single instrument to completely capture each construct. This represented a limitation of path analysis, where constructs were often measured using just one instrument. In addressing this issue, Kline (2005) said:

> One way around the aforementioned problems of measuring a construct with a single indicator is to use multiple indicators. Scores across a set of measures tend to be more reliable and valid than scores on any individual measure. Also, multiple indicators may each assess a somewhat different facet of the construct, which enhances score validity (p. 71).

To accomplish this, indices of psychosocial functioning and depressive symptom severity were created instead of relying on just one instrument to represent each construct.

With the constructs measured, the path analysis was set up to test the hypothesis that changes in depressive symptom severity mediated changes in psychosocial

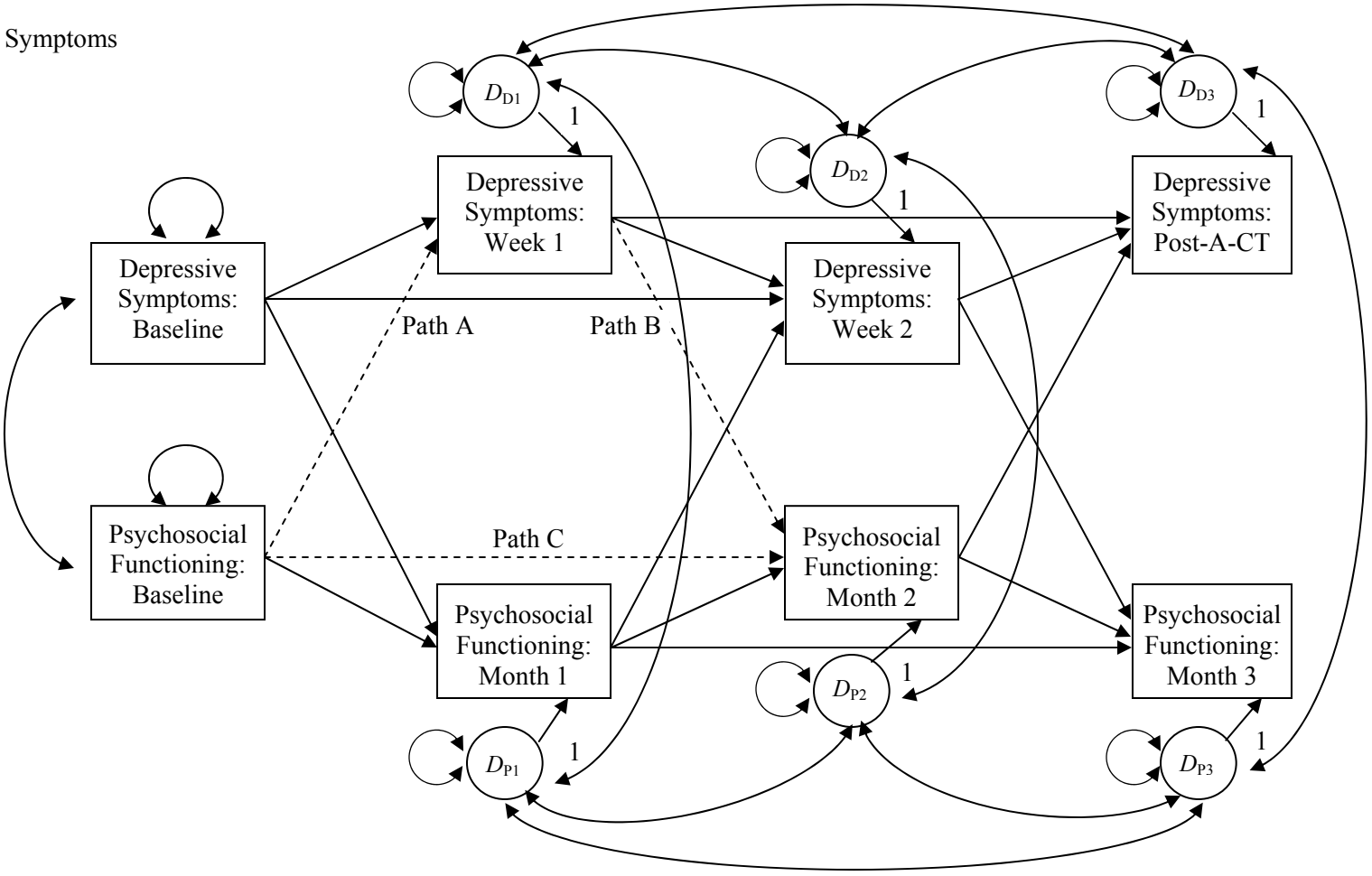
functioning during A-CT for MDD. When testing for mediating effects during treatment, Cole et al. (2003) recommended that a path analysis account for: a) levels of the independent variable (i.e., treatment vs. no treatment) measured at time one, b) the mediator measured at time one and two, and c) the dependent variable measured at time one, two, and three. With the model fit to such longitudinal data, one could then test the extent to which the indirect effects between the independent variable$_{\text{Time 1}}$, mediator $_{\text{Time 2}}$, and dependent variable $_{\text{Time 3}}$ accounted for the direct effect between the independent variable $_{\text{Time 1}}$ and dependent variable $_{\text{Time 3}}$, thus satisfying the criteria for mediating relationships (e.g., Baron et al., 1986; Wilson et al., 2002).

The path analysis in this study was based on the above recommendations by Cole et al. (2003). Since this study utilized a pre- to post-treatment design, the model was set up to look at how a mediator accounted for changes in a dependent variable compared to treatment baseline. Specifically, indices of depressive symptom severity and psychosocial functioning served as the observed, manifest variables (see Figure 4). Direct effects were then crossed-lagged and organized to account for: a) change in psychosocial functioning that occurred independently of change in depressive symptom severity, b) change in depressive symptom severity that occurred independently of change in psychosocial functioning, and c) potential mediating relationships between both constructs. As a result, this model was designed to separate variance associated with independent changes in each construct from the variance associated with mediating relationships between constructs, thereby distinguishing precedence of change from meditation.

Paths were also added to the model to control for: a) the correlation between psychosocial functioning and depressive symptom severity at treatment baseline and the beginning, middle, and end of A-CT and b) shared sources of error variance between repeated measurements. Setting the model up this way controlled for the unexplained variation in the model between constructs and across assessments.

Based on this study's conceptualization of psychosocial functioning (see p. 24), the primary aim of this study investigated the extent to which change in personal factors (depressive symptom severity) mediated change in behavior (psychosocial functioning), when people were exposed to the environmental influence of A-CT. It was hypothesized that changes in depressive symptom severity would partially mediate changes in psychosocial functioning and not vice versa. In order to be substantiated, for example, the indirect effects labeled "Path A" and "Path B" (see dashed lines in Figure 4) would need to be significant, as well as the direct effect labeled "Path C". Since a partial mediating relationship was predicted, rather than a complete mediating relationship, significant direct effects were also expected, showing depressive symptom severity and psychosocial functioning changed independently of each other.

Figure 4: Path Analysis Model that Shows Potential Mediating Relationships between Psychosocial Functioning and Depressive Symptoms



CHAPTER FIVE
Results

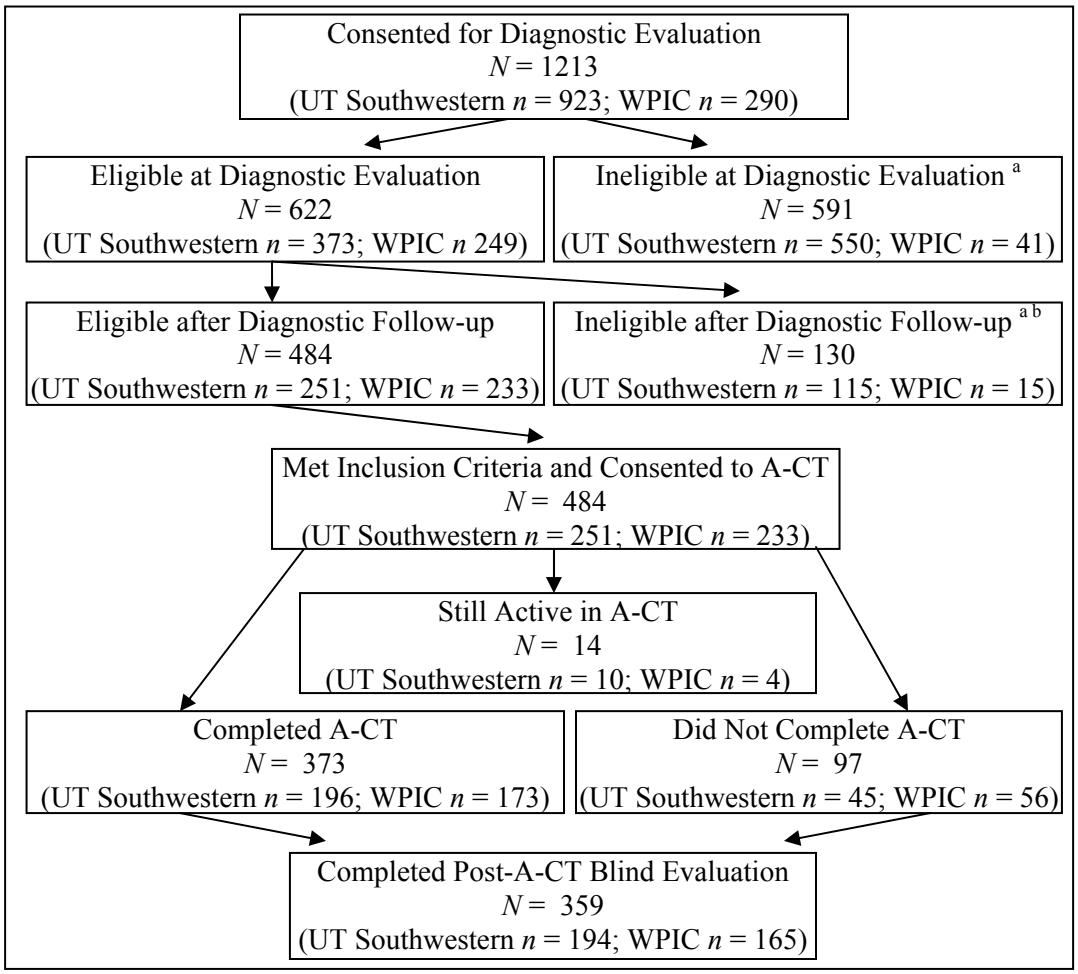
Patient Disposition

Of the 2433 patients screened by clinic staff (UT Southwestern = 923, WPIC = 1510), 1213 consented to attend an initial diagnostic evaluation (UT Southwestern = 923, WPIC = 290) (see Figure 5). Due to departmental requirements, patients at WPIC were screened for multiple studies, so there was a large discrepancy between sites in the number of patients screened and those that consented for the diagnostic evaluation. At UT Southwestern, patients were only screened for this study. Of the 1213 patients who attended an initial diagnostic evaluation, 622 (UT Southwestern n = 373; WPIC n = 249) met study criteria and were scheduled for a diagnostic follow-up, while 591 (UT Southwestern n = 550; WPIC n = 41) were excluded and referred to more appropriate treatment (see Figure 5). At the diagnostic follow-up, 484 patients (UT Southwestern n = 251; WPIC n = 233) met study eligibility requirements and signed consent to enter A-CT.

From March 1, 2000 to August 10, 2007, 373 study patients completed at least 17 (if late responder) or 15 (if early responder) sessions of A-CT, and 97 patients discontinued prematurely (see Figure 5). Fourteen patients continued to receive A-CT at the August 10, 2007, cutoff. The 470 completers and non-completers of A-CT who were no longer active in the acute-phase represented the study sample, of which 359 completed the post-A-CT blind evaluation (UT Southwestern n = 194; WPIC n = 165).

Figure 5

Patient Disposition through the End of Acute-Phase Cognitive Therapy at August 10, 2007, Cutoff



Note. A-CT = Acute-phase Cognitive Therapy; UT Southwestern = The University of Texas Southwestern Medical Center at Dallas; WPIC = Western Psychiatric Institute and Clinic at Pittsburgh.

[a] See Table 9 for reasons why patients were ineligible. [b] Eight patients await diagnostic follow-up.

Evaluators excluded a total of 721 individuals at the diagnostic evaluation and follow-up visits (see Figure 5). Table 9 summarized the rationale behind each of these exclusions. Differences across sites regarding patient exclusions reflected different referral and triage operations in use at the two university research clinics.

Table 9

Reasons for Exclusion at Diagnostic Evaluation and Follow-up

	Excluded at Diagnostic Evaluation (n = 591)		Excluded at Diagnostic Follow-up (n = 130)		
Reasons for Exclusion	UT Southwestern	WPIC	UT Southwestern	WPIC	Total
Not MDD	84	2	4	0	90
Not recurrent MDD	130	3	14	0	147
HRSD-17 < 14	31	11	18	11	71
Comorbid psychiatric disorder	80	9	24	0	113
Current alcohol/drug dependence	30	1	5	0	36
Could not comply with protocol	29	0	25	2	56
Suicidal ideation with intent	6	0	0	0	6
Previous non-response to Fluoxetine	8	0	0	0	8
Unwilling to take Fluoxetine	11	0	1	0	12
Preferred alternative treatment	19	0	11	0	30
Contraindicated medical condition	18	0	8	0	26
Contraindicated medication	22	0	3	0	25
Lost to follow-up	59	9	0	0	68
Refused further contact	15	6	2	1	24
Other [a]	8	0	0	1	9
Total	550	41	115	15	721

Note. MDD = Major Depressive Disorder; HRSD-17 = Hamilton Rating Scale for Depression-17; UT Southwestern = The University of Texas Southwestern Medical Center at Dallas; WPIC = Western Psychiatric Institute and Clinic at Pittsburgh.

[a] Other consisted of 1 patient who was a registered sex offender, 1 patient on probation, 2 patients pending/active litigation, and 5 unknowns.

Cognitive Therapist Competence

The competence of the 15 therapists in providing A-CT to study patients was examined quarterly with the Cognitive Therapy Scale (CTS; Young et al., 1980). From a total of 139 sessions that were rated, only 19 (14%) sessions, spread across eight therapists, fell below 40. Overall, and competency scores averaged 45.4 (*SD* = 5.9), with each therapist's average falling above 40 (see Appendix I for average therapist CTS). CTS scores were not related to A-CT outcomes (A. Minhajudin, personal communication, August 6, 2008).

Demographic and Clinical Characteristics of Study Patients

As seen in Table 10, study patients were most likely middle aged (*M* = 42.20, *SD* = 11.97), female (67.1%), and Caucasian (81.8%). A majority of patients were also single (57.6%) and employed either full or part time (55.2%). Based on instrument cut scores, mean total scores on measures of depressive symptom severity (e.g., BDI *M* = 28.86; HRSD-17 *M* = 21.19; IDS-SR *M* = 39.18) and psychosocial functioning (e.g., SAS-SR *M* = 2.59) showed the sample presented with moderate to severe impairment. Most patients reported having more than three major depressive episodes (77.2%) and lifetime comorbidity of other DSM-IV Axis I disorders (77.4%). Only 25 (5.3%) patients met DSM-IV criteria for current MDD and Dysthymia (i.e., double depression), and about a third of the sample met criteria for Research Diagnostic Criteria endogenous (Spitzer, Endicott, & Robins, 1975) and DSM-IV melancholic (American Psychiatric

Association, 1994) depression. Additionally, 58.7% of the sample endorsed family histories of familial depressive (i.e., patient and first degree relative were depressed) and depressive spectrum disease (i.e., first degree family members showed symptoms of alcohol abuse and/or antisocial behavior with or without depression) (Winokur, 1979).

Table 10

Baseline Demographic and Clinical Characteristics for Study Sample

Characteristics	Total Sample ($N = 470$) [a]
Age, mean ± SD, yrs	42.20 ± 11.97
Gender, No. (%) female	314 (67.1%)
Race, No. (%)	
American Indian/Alaska Native	2 (0.4%)
Asian/Pacific Islander	8 (1.7%)
Black or African American	48 (10.3%)
Hispanic/Latino	23 (4.9%)
White	383 (81.8%)
Other	6 (0.9%)
Marital Status, No. (%)	
Single	266 (57.6%)
Partnered	196 (42.4%)
Education, mean ± SD, yrs	15.03 ± 2.91
Employment outside home, No. (%)	
Full time	196 (43.8%)
Part time	51 (11.4%)
Homemaker/caregiver	27 (6.0%)
Student	21 (4.7%)
Retired	13 (2.9%)
Unemployed	96 (21.4%)
Other	44 (9.7%)
RIFT, mean ± SD	14.22 ± 2.74
SAS-SR, mean ± SD	2.59 ± 0.45
BDI, mean ± SD	28.86 ± 9.01
IDS-SR, mean ± SD	39.18 ± 10.37
HRSD-17, mean ± SD	
At initial Diagnostic Evaluation	21.19 ± 4.14
At Diagnostic Follow-up	20.30 ± 3.97
Age at onset, mean ± SD, yrs	21.16 ± 10.70
Length of current episode, mean ± SD, months	23.76 ± 39.80

Length of illness, mean ± SD, yrs	20.52 ± 11.64
No. of episodes, No. (%)	
Two episodes	99 (22.8%)
Three or more episodes	336 (77.2%)
Comorbid *DSM-IV* diagnoses, No. (%)	
Current	211 (44.9%)
Lifetime	364 (77.4%)
Current double depression, No. (%)	25 (5.3%)
Depressive subtype, No. (%)	
RDC primary depression	75 (16.0%)
RDC endogenous, definite	171 (36.4%)
DSM-IV melancholia	164 (34.9%)
DSM-III melancholia	66 (14.1%)
Family history, No. (%)	
Familial depressive disease	127 (27.0%)
Depressive spectrum disease	149 (31.7%)
Sporadic depressive disease	76 (16.2%)
Bipolar family	26 (5.5%)
Unknown	89 (18.9%)

Note. BDI = Beck Depression Inventory; HRSD = Hamilton Rating Scale for Depression-17; IDS-SR = Inventory for Depressive Symptomatology – Self-Report; RDC = Research Diagnostic Criteria; RIFT = Longitudinal Interval Follow-up Evaluation – Range of Impaired Functioning Tool; SAS-SR = Social Adjustment Scale-Self Report; UT Southwestern = University of Texas Southwestern Medical Center at Dallas; WPIC = Western Psychiatric Institute and Clinic at Pittsburgh.

[a] *n* reduced to 469 for HRSD-17 at diagnostic follow-up and DSM-III melancholia, 468 for gender, race, and RDC primary depression, 466 for age, 462 for marital status, 452 for education, 442 for initial IDS-SR, 441 for initial BDI, 435 for number of episodes, 429 for initial SAS-SR, and 352 for initial RIFT.

Preliminary Aim 1: What is the psychometric quality of the indices of psychosocial functioning and depressive symptom severity?

With prior evidence supporting the inclusion of RIFT and SAS-SR data on the index of psychosocial functioning (see Appendix G), this preliminary aim focused on evaluating the internal consistency reliability, convergent and discriminant validity, and sensitivity to change of the indices of psychosocial functioning and depressive symptom severity. While not perfect, the indices demonstrated adequate psychometric quality, justifying their inclusion in the primary aim. Evidence supporting this conclusion was described in detail below.

Do the indices of psychosocial functioning and depressive symptom severity have acceptable internal consistency reliability at treatment baseline and the beginning, middle, and end of A-CT?

While coefficients varied (see Table 11), the indices of psychosocial functioning and depressive symptom severity demonstrated adequate internal consistency reliability. For the two-item index of psychosocial functioning, the median internal consistency reliability coefficient equaled 0.65 (range = 0.49 to 0.80). As seen in Table 11, coefficients for this index were somewhat low during the month of diagnostic evaluations (alpha = 0.49) and the first month of A-CT (alpha = 0.57). Index scores from these time points were included in analyses because: a) internal consistency reliability coefficients were likely underestimated due to the low number of items on the index (Cortina, 1993; Schmitt, 1996); and b) these two scores showed convergent validity with other commonly

used measures of psychosocial functioning. For example, the index of psychosocial functioning significantly correlated with the DYS (r = -0.34), GAF (r = -0.38), and Q-LES-Q (r = -0.63) at treatment baseline and the GAF (r = -0.38) and DYS (r = -0.38) during the first month of A-CT, with p = 0.01 (The Q-LES-Q was not administered during the first month of A-CT). On the three-item index of depressive symptom severity, the median internal consistency reliability coefficient equaled 0.88 (range = 0.81 to 0.95).

Table 11

Internal Consistency Reliability Coefficients for the Indices of Psychosocial Functioning and Depressive Symptom Severity at Treatment Baseline and the Beginning, Middle, and End of Acute-Phase Cognitive Therapy

		Internal Consistency Reliability Coefficients [a]			
Index	# of Items	Diagnostic Evaluation	Week 1 of A-CT	Week 7 of A-CT	Post-A-CT Blind Evaluation
Psychosocial Functioning [b]	2	0.49	0.57	0.69	0.80
Depressive Symptom Severity	3	0.81	0.88	0.95	0.95

[a] Internal consistency reliability coefficients were calculated using Cronbach's Alpha and based on standardized items. [b] The index of psychosocial functioning included the RIFT, which measured the construct longitudinally across monthly intervals that coincided with the month of diagnostic evaluations and the first, second, and third month of A-CT.

Does the index of psychosocial functioning show convergent validity by significantly correlating with other commonly used measures of psychosocial functioning (e.g., DYS, GAF, and Q-LES-Q) at treatment baseline and the beginning, middle, and end of A-CT?

As mentioned in the previous section, the index of psychosocial functioning showed convergent validity by being significantly correlated with other commonly used measures of psychosocial functioning (see Table 12). Specifically, across time points, the index of psychosocial functioning significantly correlated with the DYS (*r* ranged from -0.34 to -0.54), GAF (*r* ranged from -0.38 to - 0.69), and Q-LES-Q (*r* ranged from -0.63 to -0.82), with $p = 0.01$. Correlations were negative because higher scores indicated improved functioning on the DYS, GAF, and Q-LES-Q and deteriorated functioning on the index of psychosocial functioning.

Table 12

Correlation Coefficients between the Index of Psychosocial Functioning and Total Scores from the DYS, GAF, and Q-LES-Q

Index of Psychosocial Functioning	Correlation Coefficient		
	DYS	GAF	Q-LES-Q [a]
Month of diagnostic evaluations	-0.34*	-0.38*	-0.63*
First month of A-CT	-0.44*	-0.41*	
Second month of A-CT	-0.43*	-0.61*	
Third month of A-CT	-0.54*	-0.69*	-0.82*

Note. DYS = Dyadic Adjustment Scale; GAF = Global Assessment of Functioning Scale; Q-LES-Q = Quality of Life Enjoyment and Satisfaction Questionnaire.

[a] The Q-LES-Q was only administered during the diagnostic evaluation and post-A-CT blind evaluation.

* $p = 0.01$

Does the index of psychosocial functioning show discriminant validity by not being highly correlated (i.e., r < .50; Cohen, 1988) with the BDI, HRSD-17, IDS-SR, and index of depressive symptom severity at treatment baseline and the beginning, middle, and end of A-CT?

The index of psychosocial functioning did not show discriminant validity due to its high correlations with the BDI, HRSD-17, IDS-SR, and index of depressive symptom severity. As seen in Table 13, correlations between the index of psychosocial functioning and measures of depressive symptom severity ranged from 0.39 to 0.77. With all but two exceptions (i.e., HRSD-17 at month of diagnostic evaluations and the first month of A-CT), correlations were large in size, exceeding the 0.50 cutoff suggested by Cohen (1988).

Table 13

Correlation Coefficients between the Index of Psychosocial Functioning and the Index of Depressive Symptom Severity, BDI, HRSD-17, and IDS-SR

Index of Psychosocial Functioning	Correlation Coefficient			
	BDI	HRSD-17	IDS-SR	Index of Depressive Symptom Severity
Month of diagnostic evaluations	0.59*	0.39*	0.58*	0.61*
First month of A-CT	0.63*	0.48*	0.59*	0.62*
Second month of A-CT	0.68*	0.65*	0.69*	0.69*
Third month of A-CT	0.76*	0.72*	0.77*	0.77*

Note. BDI = Beck Depression Inventory; HRSD-17 = Hamilton Rating Scale for Depression-17; IDS-SR = Inventory for Depressive Symptomatology – Self-Report.

* $p = 0.01$

Do the indices of psychosocial functioning and depressive symptom severity show sensitivity to changes during A-CT by having large pre- to post-A-CT standardized mean gain effect sizes (i.e., $d \geq 0.80$; Cohen, 1988)?

Results from this analysis showed the two indices in this study were sensitive to pre- to post-A-CT changes. Specifically, large standardized mean gain effect sizes of 1.54 and 2.37 represented pre- to post-A-CT changes on the indices of psychosocial functioning and depressive symptom severity, respectively (see Table 14). These effect sizes were similar to those reported in RCTs investigating the impact of A-CT on psychosocial functioning (i.e., $d = 1.39$; see Chapter Three) and depressive symptom

severity (e.g., d = 2.40 in Elkin et al., 1989; d = 1.46 in Thompson et al., 1987). When comparing effect sizes, depressive symptom severity appeared to change more than psychosocial functioning, with most change occurring between week one and seven of A-CT (see Table 14). As seen in Figure 6, change slopes were close to parallel except between week one and seven of A-CT, when depressive symptom severity improved more than psychosocial functioning.

Table 14

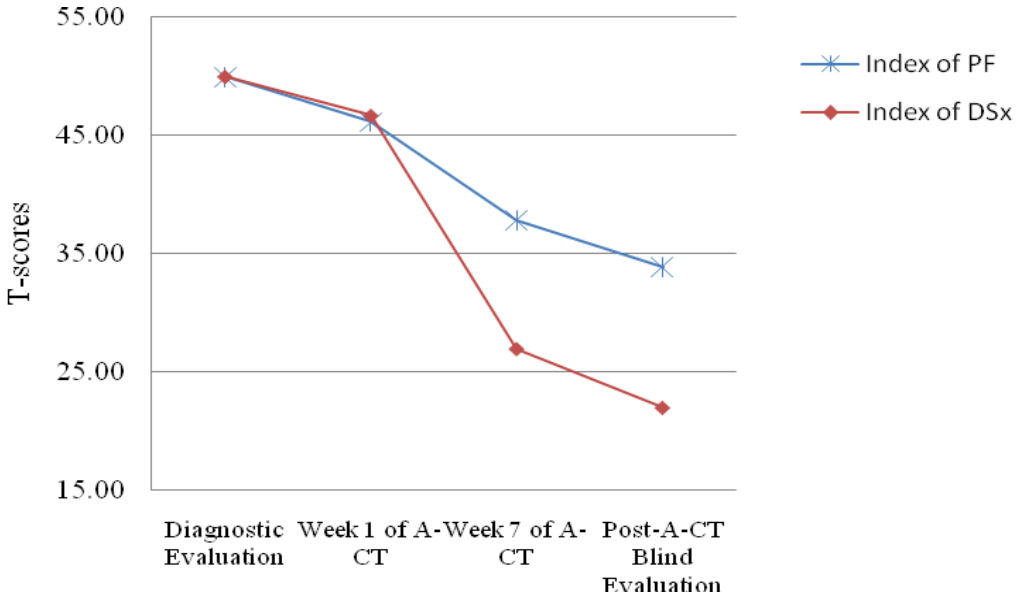
Standardized Mean Gain Effect Sizes for Indices of Psychosocial Functioning and Depressive Symptom Severity Showing Change from Treatment Baseline

	Standardized Mean Gain Effect Size (d)		
Index	Baseline to Week 1 of A-CT	Baseline to Week 7 of A-CT	Baseline to Post-A-CT Blind Evaluation
Psychosocial Functioning [a]	0.38	1.16	1.54
Depressive Symptom Severity	0.31	1.94	2.37

[a] The index of psychosocial functioning included the RIFT, which measured the construct longitudinally across monthly intervals that coincided with the month of diagnostic evaluations and the first, second, and third month of A-CT.

Figure 6

Change Slopes for Indices of Psychosocial Functioning and Depressive Symptom Severity during Acute-Phase Cognitive Therapy



Note. Index of PF = Index of Psychosocial Functioning; Index of DSx = Index of Depressive Symptom Severity.

Summary of Preliminary Aim 1

This preliminary aim focused on evaluating the psychometric quality of the indices of psychosocial functioning and depressive symptom severity. Results replicated previous literature (e.g., Vittengl et al., 2004), showing the index of depressive symptom severity had high internal consistency reliability and sensitivity to pre- to post A-CT changes. Also, results found the index of psychosocial functioning had acceptable

internal consistency reliability and convergent validity with other commonly used measures of psychosocial functioning (e.g., DYS, GAF, and Q-LES-Q).

Of note, the index of psychosocial functioning did not show discriminant validity due to its high correlation with measures of depressive symptom severity. In fact, at the post-A-CT blind evaluation, the indices of psychosocial functioning and depressive symptom severity shared over half of their variability in scores (r = 0.77, R^2 = 0.59). This finding, however, did not warrant the removal of the index of psychosocial functioning from the primary aim, since it replicated previous research (see p. 37 for a review of this literature) in illustrating a limitation in the field more so than a limitation in the measurements in this study. Therefore, taken as a whole, this preliminary aim supported the psychometric quality of the indices of psychosocial functioning and depressive symptom severity, substantiating their use in the primary aim.

Preliminary Aim 2: Do baseline demographic and clinical characteristics differ at UT Southwestern compared to WPIC?

Overall, compared to patients from WPIC, patients from UT Southwestern were more likely to be from racial minorities, related to first degree relatives with a history of psychopathology, and diagnosed with depression of greater severity and chronicity. As seen in Table 15, a greater percentage of patients from UT Southwestern endorsed depressive spectrum disease (χ^2 [5, N = 470] = 19.42, p = 0.002) and a Hispanic/Latino race (χ^2 [5, N = 468] = 28.63, p < 0.001). In addition, evaluators diagnosed patients from UT Southwestern as more chronically depressed, with earlier ages of onset (t [444] = -

3.37, $p = 0.001$), longer current major depressive episodes (t [338] = 5.41, $p < 0.001$), and longer durations of depressive illness (t [468] = 3.29, $p = 0.001$). Patients from UT Southwestern also self-reported greater baseline depressive symptom severity, with total scores on the BDI reflecting severe depression for UT Southwestern patients and moderate to severe depression for WPIC patients (t [439] = 4.02, $p < 0.001$). While statistically significant, site differences on the SAS-SR, IDS-SR, and HRSD-17 did not reflect clinically meaningful differences, as total scores from both sites reflected severe impairment.

Table 15

Baseline Demographic and Clinical Characteristics for Study Sample by Study Site

Characteristics	UT Southwestern ($n = 241$) [a]	WPIC ($n = 229$) [b]	Statistics
Age, mean ± SD, yrs	42.18 ± 11.79	42.23 ± 12.18	$p = 0.96$ [c]
Gender, No. (%) female	169 (70.1%)	145 (63.9%)	$p = 0.17$ [d]
Race, No. (%)			
American Indian/Alaska Native	2 (0.8%)	0 (0%)	
Asian	6 (2.4%)	2 (0.8%)	
Black or African American	25 (10.4%)	23 (10.1%)	$p < 0.001$ [e]
Hispanic/Latino	23 (9.5%)	0 (0%)	
White	182 (75.5%)	201 (88.5%)	
Other	3 (1.2%)	1 (0.4%)	
Marital Status, No. (%)			
Single	131 (54.4%)	135 (61.1%)	$p = 0.45$ [e]
Partnered	110 (45.6%)	86 (38.9%)	
Education, mean ± SD, yrs	15.15 ± 2.74	14.90 ± 3.10	$p = 0.37$ [c]
Employment outside home, No. (%)			
Full time	103 (42.7%)	93 (44.9%)	
Part time	30 (12.4%)	21 (10.1%)	
Homemaker/caregiver	13 (5.4%)	14 (6.8%)	
Student	10 (4.1%)	11 (5.3%)	$p = 0.19$ [e]
Retired	8 (3.3%)	5 (2.4%)	
Other	27 (11.2%)	17 (8.2%)	
Unemployed	50 (20.7%)	46 (22.2%)	

RIFT, mean ± SD	14.17 ± 2.92	14.28 ± 2.52	$p = 0.70$ [c]
SAS-SR, mean ± SD	2.66 ± 0.46	2.51 ± 0.42	$p = 0.001$ [c]
BDI, mean ± SD	30.41 ± 8.41	27.00 ± 9.37	$p < 0.001$ [c]
IDS-SR, mean ± SD	40.56 ± 10.03	37.53 ± 10.56	$p = 0.002$ [c]
HRSD-17, mean ± SD			
At initial Diagnostic Evaluation	20.52 ± 4.09	21.88 ± 4.08	$p < 0.001$ [c]
At Diagnostic Follow-up	19.38 ± 3.66	21.29 ± 4.06	$p < 0.001$ [c]
Age at onset, mean ± SD, yrs	19.55 ± 9.58	22.86 ± 11.54	$p = 0.001$ [c]
Length of current episode, mean ± SD, mo	33.02 ± 49.42	14.02 ± 22.41	$p < 0.001$ [c]
Length of illness, mean ± SD, yrs	22.22 ± 11.71	18.73 ± 11.31	$p = 0.001$ [c]
No. of episodes, No. (%)			
Two episodes	58 (24.3%)	41 (20.9%)	$p = 0.42$ [d]
Three or more episodes	181 (75.7%)	155 (79.1%)	
Comorbid *DSM-IV* diagnoses, No. (%)			
Current	106 (44.0%)	105 (46.1%)	$p = 0.71$ [d]
Lifetime	178 (73.9%)	185 (81.1%)	$p = 0.06$ [d]
Current double depression, No. (%)	9 (3.7%)	16 (7.0%)	$p = 0.15$ [d]
Depressive subtype, No. (%)			
RDC primary depression	197 (82.1%)	192 (84.6%)	$p = 0.06$ [e]
RDC endogenous, definite	87 (36.1%)	83 (36.4%)	$p = 0.08$ [e]
DSM-IV melancholia	80 (33.2%)	83 (36.4%)	$p = 0.14$ [e]
DSM-III melancholia	27 (11.3%)	38 (16.7%)	$p = 0.21$ [e]
Family history, No. (%)			
Familial depressive disease	65 (27.0%)	62 (27.1%)	
Depressive spectrum disease	87 (36.1%)	62 (27.1%)	
Sporadic depressive disease	41 (17.0%)	35 (15.3%)	$p = 0.002$ [e]
Bipolar family	18 (7.5%)	8 (3.5%)	
Unknown	29 (12.0%)	60 (26.2%)	

Note. BDI = Beck Depression Inventory; HRSD = Hamilton Rating Scale for Depression-17; IDS-SR = Inventory for Depressive Symptomatology – Self-Report; RDC = Research Diagnostic Criteria; RIFT = Longitudinal Interval Follow-up Evaluation – Range of Impaired Functioning Tool; SAS-SR = Social Adjustment Scale-Self Report; UT Southwestern = University of Texas Southwestern Medical Center at Dallas; WPIC = Western Psychiatric Institute and Clinic at Pittsburgh.

[a] *n* reduced to 240 for RDC primary depression, DSM-III melancholia, initial BDI, and initial IDS-SR, 239 for number of episodes, 231 for initial SAS-SR, and 189 for initial RIFT. [b] *n* reduced to 228 for current comorbidity, past comorbidity, double depression, RDC endogenous, DSM-IV melancholia, DSM-III melancholia, and HRSD-17 at diagnostic follow-up, 227 for gender, race, and RDC primary depression, 225 for age, 221 for marital status, 211 for education, 207 for employment status, 202 for initial IDS-SR, 201 for initial BDI, 198 for initial SAS-SR, 196 for number of episodes, and 163 for initial RIFT. [c] Probability value comes from *t*-test for independent means. [d] Probability value comes from Fisher's exact test. [e] Probability value comes from Pearson Chi-Square.

Preliminary Aim 3: Do early and late responders to acute-phase Cognitive Therapy differ across baseline demographic and clinical characteristics?

According to this preliminary aim, early responders were more likely to have full time employment, less depressive symptom severity, and shorter major depressive episodes. As seen in Table 16, early and late responders to A-CT endorsed significantly different types of employment (χ^2 [9, N = 448] = 24.02, p = 0.004), with early responders having higher percentages of full-time employment and lower percentages of unemployment. Also, evaluators described the current major depressive episode of early responders as significantly shorter in duration than that of late responders (t [454] = -3.80, $p < 0.001$). At the initial diagnostic evaluation, early responders self-reported lower depressive symptom severity, with total scores on the BDI showing moderate to severe

depression for early responders and severe depression for late responders (t [439] = 3.49, p = 0.001). As before, statistically significantly differences between early and late responders on total scores from the RIFT, SAS-SR, IDS-SR, and HRSD-17 did not reflect clinically meaningful differences, as total scores irrespective of response type reflected severe impairment.

Table 16

Baseline Demographic and Clinical Characteristics for Study Sample by Pattern of Treatment Response (Early vs. Late)

Characteristics	Early Responders (n = 178) [a]	Late Responders (n = 292) [b]	Statistics
Age, mean ± SD, yrs	42.45 ± 12.15	42.05 ± 11.88	p = 0.73 [c]
Sex, No. (%) female	118 (66.7%)	196 (67.4%)	p = 0.92 [d]
Race, No. (%)			
American Indian/Alaska Native	1 (0.6%)	1 (0.3%)	
Asian	3 (1.7%)	5 (1.7%)	
Black or African American	14 (7.9%)	34 (11.7%)	p < 0.47 [e]
Hispanic/Latino	8 (4.5%)	15 (5.2%)	
White	151 (85.3%)	232 (79.7%)	
Other	0 (0%)	4 (1.4%)	
Marital Status, No. (%)			
Single	102 (58.3%)	164 (57.1%)	p = 0.62 [e]
Partnered	73 (41.7%)	123 (42.9%)	
Education, mean ± SD, yrs	15.10 ± 2.88	14.99 ± 2.94	p = 0.71 [c]
Employment outside home, No. (%)			
Full time	89 (52.7%)	107 (38.4%)	
Part time	15 (8.9%)	36 (12.9%)	
Homemaker/caregiver	9 (5.3%)	18 (6.5%)	
Student	4 (2.4%)	17 (6.1%)	p = 0.004 [e]
Retired	6 (3.6%)	7 (2.5%)	
Other	21 (12.4%)	23 (8.2%)	
Unemployed	25 (14.8%)	71 (25.4%)	
RIFT, mean ± SD	13.69 ± 2.62	14.67 ± 2.76	p = 0.001 [c]
SAS-SR, mean ± SD	2.51 ± 0.43	2.64 ± 0.45	p = 0.002 [c]
BDI, mean ± SD	26.98 ± 8.78	30.02 ± 8.97	p = 0.001 [c]

IDS-SR, mean ± SD	36.05 ± 10.30	41.08 ± 9.97	$p < 0.001$ [c]
HRSD-17, mean ± SD			
At initial Diagnostic Evaluation	20.28 ± 3.96	21.74 ± 4.15	$p < 0.001$ [c]
At Diagnostic Follow-up	19.46 ± 3.83	20.82 ± 3.97	$p < 0.001$ [c]
Age at onset, mean ± SD, yrs	21.17 ± 10.65	21.15 ± 10.74	$p = 0.98$ [c]
Length of current episode, mean ± SD, mo	16.13 ± 23.38	28.42 ± 46.51	$p < 0.001$ [c]
Length of illness, mean ± SD, yrs	20.63 ± 12.03	20.45 ± 11.41	$p = 0.87$ [c]
No. of episodes, No. (%)			
Two episodes	30 (18.3%)	69 (25.5%)	$p = 0.10$ [d]
Three or more episodes	134 (81.7%)	202 (74.5%)	
Comorbid *DSM-IV* diagnoses, No. (%)			
Current	74 (41.8%)	137 (46.9%)	$p = 0.29$ [d]
Lifetime	129 (72.9%)	234 (80.1%)	$p = 0.09$ [d]
Current double depression, No. (%)	10 (5.6%)	15 (5.1%)	$p = 0.83$ [d]
Depressive subtype, No. (%)			
RDC primary depression	154 (87.0%)	235 (81.0%)	$p = 0.24$ [e]
RDC endogenous, definite	62 (35.0%)	108 (37.0%)	$p = 0.66$ [e]
DSM-IV melancholia	65 (36.7%)	98 (33.6%)	$p = 0.44$ [e]
DSM-III melancholia	25 (14.1%)	40 (13.7%)	$p = 0.33$ [e]
Family history, No. (%)			
Familial depressive disease	44 (24.7%)	83 (28.4%)	
Depressive spectrum disease	55 (30.9%)	94 (32.2%)	
Sporadic depressive disease	37 (20.8%)	39 (13.4%)	$p = 0.33$ [e]
Bipolar family	7 (3.9%)	19 (6.5%)	
Unknown	34 (19.1%)	55 (18.8%)	
Site Location, No. (%)			
UT Southwestern (Dallas)	88	151	$p = 0.69$ [d]
WPIC (Pittsburgh)	77	120	

Note. BDI = Beck Depression Inventory; HRSD = Hamilton Rating Scale for Depression-17; IDS-SR = Inventory for Depressive Symptomatology – Self-Report; RDC = Research Diagnostic Criteria; RIFT = Longitudinal Interval Follow-up Evaluation – Range of Impaired Functioning Tool; SAS-SR = Social Adjustment Scale-Self Report; UT Southwestern = University of Texas Southwestern Medical Center at Dallas; WPIC = Western Psychiatric Institute and Clinic at Pittsburgh.

[a] *n* reduced to 177 for gender, race, current comorbidity, past comorbidity, double depression, RDC endogenous, DSM-IV melancholia, DSM-III melancholia, RDC primary depression, and HRSD-17 at diagnostic follow-up, 176 for age, 175 for marital status, 171 for education, 169 for employment status and initial BDI, 167 for initial IDS-SR, 164 for number of episodes and site, and 162 for initial RIFT and SAS-SR. [b] *n* reduced to 291 for gender, race, and DSM-III melancholia, 290 for age and RDC primary depression, 287 for marital status, 281 for education, 279 for employment, 275 for initial IDS-SR, 272 for initial BDI, 271 for number of episodes and site, 267 for initial SAS-SR, and 190 for initial RIFT. [c] Probability value comes from *t*-test for independent means. [d] Probability value comes from Fisher's exact test. [e] Probability value comes from Pearson Chi-Square.

Preliminary Aim 4: Does depressive symptom severity significantly reduce after exposure to A-CT compared to pre-treatment baseline?

For this preliminary aim, analyses were first run to determine how much variability in pre- to post-A-CT changes in depressive symptom severity were attributable to exposure to A-CT vs. some pre-existing condition (i.e., site difference or early vs. late response) or exposure to other treatments. The proposed second step, which would have examined the relative impact of each confounding variable, was not needed because only one variable (i.e., early vs. late responder to A-CT) influenced pre- to post-A-CT changes in depressive symptom severity. Once the variability associated with exposure to A-CT was better isolated by controlling for an early vs. late response to A-CT, it was predicted

that the index of depressive symptom severity would significantly lower from pre- to post-A-CT. Results confirmed this hypothesis.

Investigation of the Impact of Potentially Confounding Variables on Reductions on the Index Depressive Symptom Severity

As a first step, the impact of potentially confounding variables (i.e., study site, early vs. late response to A-CT, and presence or absence of additional treatment) on pre- to post-A-CT changes in depressive symptom severity was examined. While study site (F [3, 691] = 0.18, p = 0.91) did not significantly impact pre- to post-A-CT changes in depressive symptom severity, an early vs. late response pattern to A-CT did (F [3, 701] = 11.78, $p < 0.001$) (see Table 17). Post hoc comparisons showed early responders to A-CT endorsed significantly less depressive symptom severity than late responders at diagnostic evaluation, week one and seven of A-CT, and the post-A-CT blind evaluation (see Figure 7). Post hoc comparisons, using the Bonferroni method, were significant at the p = 0.001 level and based on scores of the index of depressive symptom severity.

Table 17

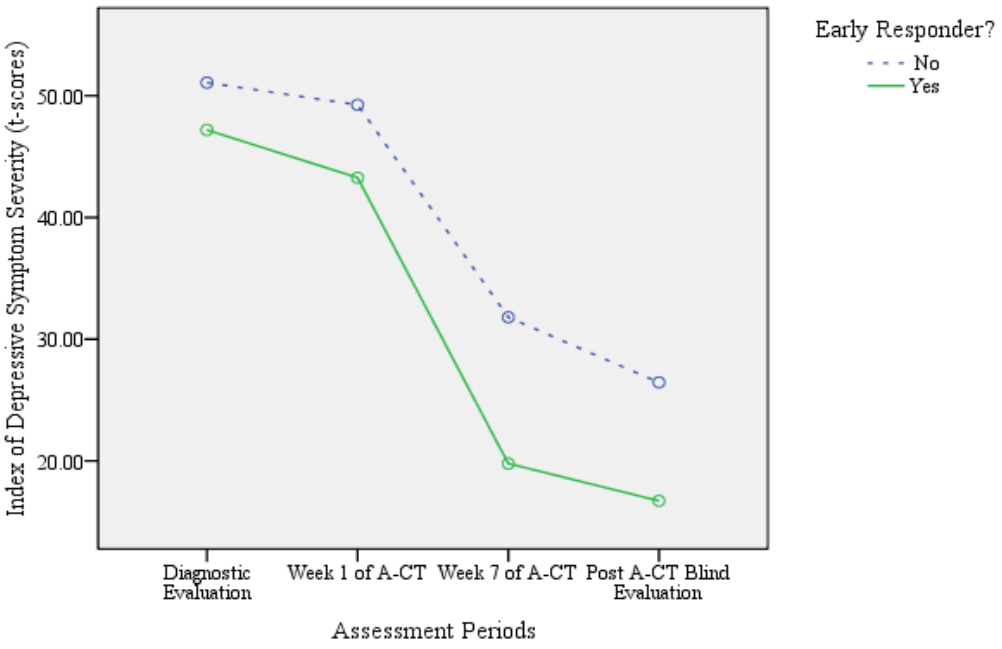
Effect of Time in Acute-Phase Cognitive Therapy on the Index of Depressive Symptom Severity by an Early vs. Late Response

	Type III Tests of Fixed Effects				
Source	Numerator *df*	Denominator *df*	*F*-value	*p*	R^2
Intercept	1	1437	15203.26	0.000	
Visit	3	701	582.96	0.000	0.46
Early vs. Late Response	1	1437	199.08	0.000	0.12
Visit*Early vs. Late Response	3	701	11.78	0.000	0.02

Note. Visit = the index of depressive symptom severity measured across repeated measurements (e.g., diagnostic evaluation, week one and seven of A-CT, and post-A-CT blind evaluation).

Figure 7

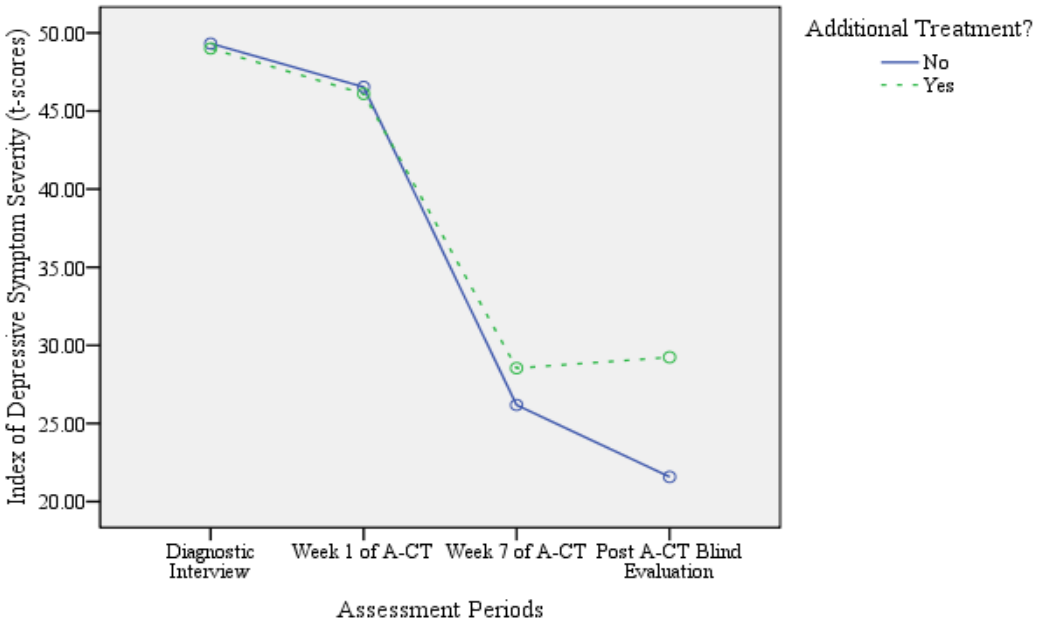
Pre- to Post-Acute-Phase Cognitive Therapy Reductions on the Index of Depressive Symptom Severity by an Early vs. Late Response



While not statistically significant, a trend was noted for the influence of additional treatment on pre- to post-A-CT changes in depressive symptom severity (F [3, 719] = 2.19, p = 0.09). As seen in Figure 8, patients with additional treatment endorsed somewhat higher depressive symptom severity at the post-A-CT blind evaluation than patients without additional treatment, with means on the index of depressive symptom severity equaling 30.25 (SD = 16.95) and 21.52 (SD = 13.52), respectively.

Figure 8

Pre- to Post-Acute-Phase Cognitive Therapy Reductions on the Index of Depressive Symptom Severity across Presence or Absence of Additional Treatment



Pre- to Post-A-CT Reductions in Depressive Symptom Severity: Controlling for Confounding Variables

Since only one between-group variable (i.e., early vs. late response to A-CT) significantly influenced pre- to post-A-CT changes in depressive symptom severity, step two was not needed. Again, in step two, a repeated-measures ANOVA would have been run to examine pre- to post-A-CT changes in depressive symptom severity, controlling for all between-group factors that significantly influenced this change in step one. This step was initially planned to understand the relative impact of each confounding variable,

since it was possible that the influence of one between-group factor was accounted for by another.

Instead, the extent to which depressive symptom severity changed from pre- to post-A-CT, controlling for the variability associated with confounding variables, could be understood by studying Table 17. As seen in Table 17, depressive symptom severity significantly lowered from pre- to post-A-CT (F [3, 701] = 582.96, $p < 0.001$), after controlling for the influence of an "early vs. late response to A-CT". Post-hoc analyses showed significant reductions ($p < 0.001$) on the index of depressive symptom severity occurred between the diagnostic evaluation ($M = 49.48$, $SE = 0.46$), week one of A-CT ($M = 46.12$, $SE = 0.53$), week seven of A-CT ($M = 26.31$, $SE = 0.62$), and the post-A-CT blind evaluation ($M = 21.57$, $SE = 0.69$).

As also seen in Table 17, the within-groups factor "time in A-CT" (i.e., index of depressive symptom severity measured at diagnostic evaluation, week one and seven of A-CT, and post-A-CT blind evaluation) accounted for 46% of the variability in reductions on the index of depressive symptom severity. On the other hand, the interaction between "time in A-CT" and an "early vs. late response to A-CT" accounted for only 2% of the variability in pre- to post-A-CT reductions on the index of depressive symptom severity. Consequently, results suggested that the confounding variable (i.e., "early vs. late response to A-CT") had little impact on pre- to post-A-CT changes in depressive symptom severity and did not need to be included in the final model.

What Percentage of Patients Responded to A-CT?

As mentioned before, treatment response was defined as both the absence of MDD and an HRSD-17 of ≤ 12. When determining treatment response, data were used from the post-A-CT blind evaluation or last available sessions of A-CT, if the patient dropped or withdrew during the acute-phase (see p. 95 for additional information on definition of treatment response). A total of 276 out of 470 patients (59%) responded to A-CT (59% at UT Southwestern and 59% at WPIC). This finding replicated what Holon et al. (2005) reported (e.g., 58%) in a RCT using a similar definition of treatment response to A-CT.

Summary of Preliminary Aim 4

Overall, results from this preliminary aim found exposure to A-CT was associated with significant reductions in depressive symptom severity. This significant reduction was substantiated even when controlling for the variability in index scores taken up by early vs. late response patterns to A-CT. Also, since response rates were similar to those reported in RCTs of A-CT, it could be inferred that the treatment response observed in this study was comparable to exposure to A-CT, rather than regression to the mean or some other confounding variable.

Preliminary Aim 5: To what extent does psychosocial functioning improve after exposure to A-CT compared to pre-treatment baseline?

As with preliminary aim four, analyses were first run to determine how much variability in pre- to post-A-CT changes in psychosocial functioning were attributable to exposure to A-CT vs. some pre-existing condition (i.e., site difference or early vs. late response) or exposure to other treatments. Again, only one potentially confounding variable (i.e., early vs. late response) was found that significantly impacted pre- to post-A-CT changes in psychosocial functioning. Once the variability associated with an early vs. late response to A-CT was controlled, it was predicted that impairment in psychosocial functioning would significantly improve from pre- to post-A-CT. Results confirmed this hypothesis.

Investigation of the Impact of Potentially Confounding Variables on Changes in Psychosocial Functioning

Analyses in this preliminary aim first determined the extent to which embedded factors in the study design (i.e., study site and early vs. late response to A-CT) and the "presence or absence of additional treatment" accounted for pre- to post-A-CT changes in psychosocial functioning. According to results, levels of impairment in psychosocial functioning, measured by the index of psychosocial functioning, did not differ across study site (F [3, 691] = 0.43, p = 0.73) or the presence or absence of additional treatment (F [3, 715] = 0.90, p = 0.44). However, early responders to A-CT reported significantly lower impairment in psychosocial functioning than late responders, measured by the

index of psychosocial functioning (F [3, 691] = 3.70, p = 0.01) (see Table 18). Specifically, post hoc comparisons showed early responders reported significantly less impairment in psychosocial functioning than late responders at treatment baseline and the first, second, and third month of A-CT (see Figure 9). All post hoc comparisons used the Bonferroni method and were significant at $p = 0.001$.

Table 18

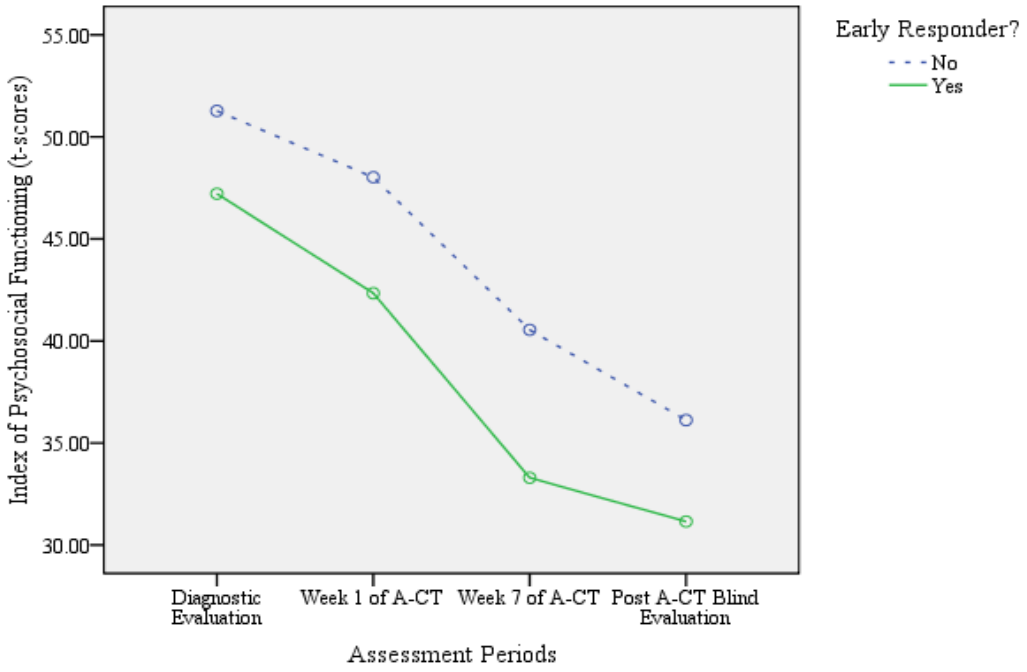
Effect of Time in Acute-Phase Cognitive Therapy on the Index of Psychosocial Functioning by an Early vs. Late Response

	Type III Tests of Fixed Effects				
Source	Numerator *df*	Denominator *df*	*F*-value	*p*	R^2
Intercept	1	1512.761	25824.22	.000	
Visit	3	691.458	197.63	.000	0.22
Early vs. Late Response	1	1512.761	113.32	.000	0.07
Visit*Early vs. Late Response	3	691.458	3.70	.012	0.01

Note. Visit = the index of psychosocial functioning measured across repeated measurements (e.g., month of diagnostic evaluations and first, second, and third month of A-CT).

Figure 9

Pre- to Post-Acute-Phase Cognitive Therapy Reductions on the Index of Psychosocial Functioning by an Early vs. Late Response



Pre- to Post-A-CT Reductions in Psychosocial Functioning: Controlling for Confounding Variables

As in preliminary aim four, step two was not needed to understand the relative impact of each confounding variable, since only one between-groups factor (i.e., "early vs. late response to A-CT") significantly influenced pre- to post-A-CT changes in psychosocial functioning. As seen in Table 18, impairment in psychosocial functioning significantly improved from pre- to post-A-CT (F [3, 691] = 197.63, $p < 0.001$), after an

"early vs. late response to A-CT" was controlled. Bonferroni post hoc comparisons found significant reductions (p = 0.001) occurred on the index of psychosocial functioning between the month of diagnostic evaluations (M = 49.60, SE = 0.48) and month one (M = 45.63, SE = 0.48), two (M = 37.42, SE = 0.53), and three of A-CT (M = 33.64, SE = 0.58).

As also seen in Table 18, the within-groups variable "time in A-CT" accounted for 22% of the variability in reduced scores on the index of psychosocial functioning across repeated measurements. Comparatively, the interaction between "time in A-CT" and an "early vs. late response to A-CT" accounted for very little variability in predicting changes in psychosocial functioning during A-CT (R^2 = 0.01). According to these results, it was concluded that an early vs. late response to A-CT did not need to be controlled when examining pre- to post-A-CT changes in psychosocial functioning.

Summary of Preliminary Aim 5

In sum, results found exposure to A-CT was associated with significant reductions in impairment in psychosocial functioning associated with MDD. This significant reduction was substantiated even when controlling for the variability taken up by early vs. late response patterns to A-CT.

Primary Aim: To what extent does change in variance unique to depressive symptom severity mediate change in variance unique to psychosocial functioning?

To address the primary aim, a structural equation model was fit with data representing depressive symptom severity (measured by the index of depressive symptom severity) and psychosocial functioning (measured by the index of psychosocial functioning) at treatment baseline and the beginning, middle, and end of A-CT. It was expected that the model would show significant, independent changes in each construct during A-CT and a partial mediating relationship, such that changes in variance unique to depressive symptom severity mediated changes in variance unique to psychosocial functioning. The path analysis model initially demonstrated poor fit with the data, but fit improved to acceptable levels after correcting specification errors, trimming non-significant paths, and building two additional paths. Results partially confirmed the hypotheses. Specifically, while independent changes in each construct were substantiated, it appeared that changes in variance unique to psychosocial functioning partially mediated changes in variance unique to depressive symptom severity, and not vice versa. In what follows, each of these finding were explained in greater detail.

Primary Aim: Model Fit

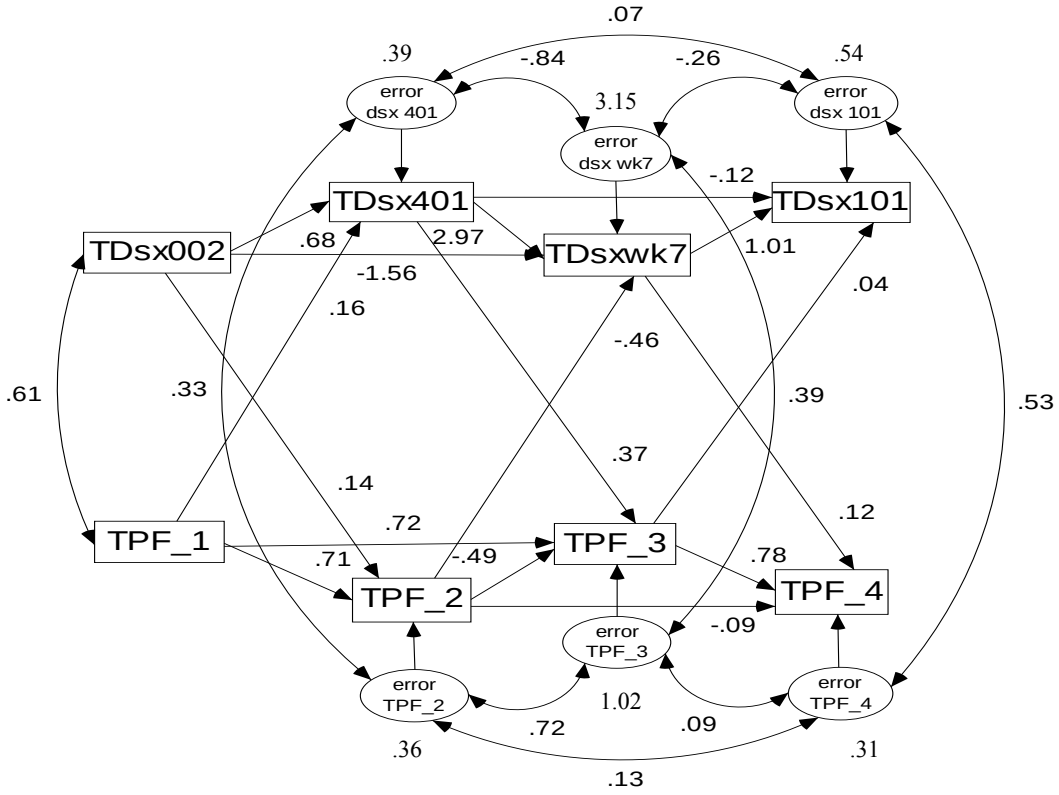
Before a structural equation model should be interpreted, it must first fit the data. The phrase "fit the data" meant that the relationships between data in the model did not significantly differ from what was expected in a normally distributed population. The model evaluated in this study was fit with data to account for: a) covariation between

depressive symptom severity and psychosocial functioning at treatment baseline and the beginning, middle, and end of A-CT, b) shared sources of error variance across repeated measurements, c) change in depressive symptom severity that occurred independently of change in psychosocial functioning, d) change in psychosocial functioning that occurred independently of change in depressive symptom severity, and e) potential mediating relationships between depressive symptom severity and psychosocial functioning (see Figure 4). It was predicted that the path analysis model would show acceptable fit across four indices (i.e., χ_m^2, RMSEA, CFI, and NFI). Results did not initially confirm this hypothesis.

For the model in Figure 10, fit indices were mixed. While the NFI and CFI equaled 1.00, which indicated optimal fit, the model chi square (χ_m^2) equaled 6.29, p = .04, which indicated poor fit. Also, the RMSEA equaled 0.07 (90% CI = 0.01 to 0.13), reflecting moderate fit (values in the 90% CI should also be below 0.10). The model in Figure 10 also contained two Heywood cases, which were defined as "parameter estimates with illogical values" (e.g., correlations > 1.0; p. 114; Kline, 2005). Since parameter estimates for disturbances represented the proportion of variability in endogenous variables that was unexplained by the model, estimates could not exceed 1.0. In Figure 10, however, disturbance estimates for the indices of depressive symptom severity at week seven of A-CT (TDsxWK7) and psychosocial functioning at month two of A-CT (TPF_3) both exceeded one, equaling 3.15 and 1.02 respectively.

Figure 10

Initial Path Analysis Model Showing Mediating Relationships between Psychosocial Functioning and Depressive Symptoms: Standardized Solution

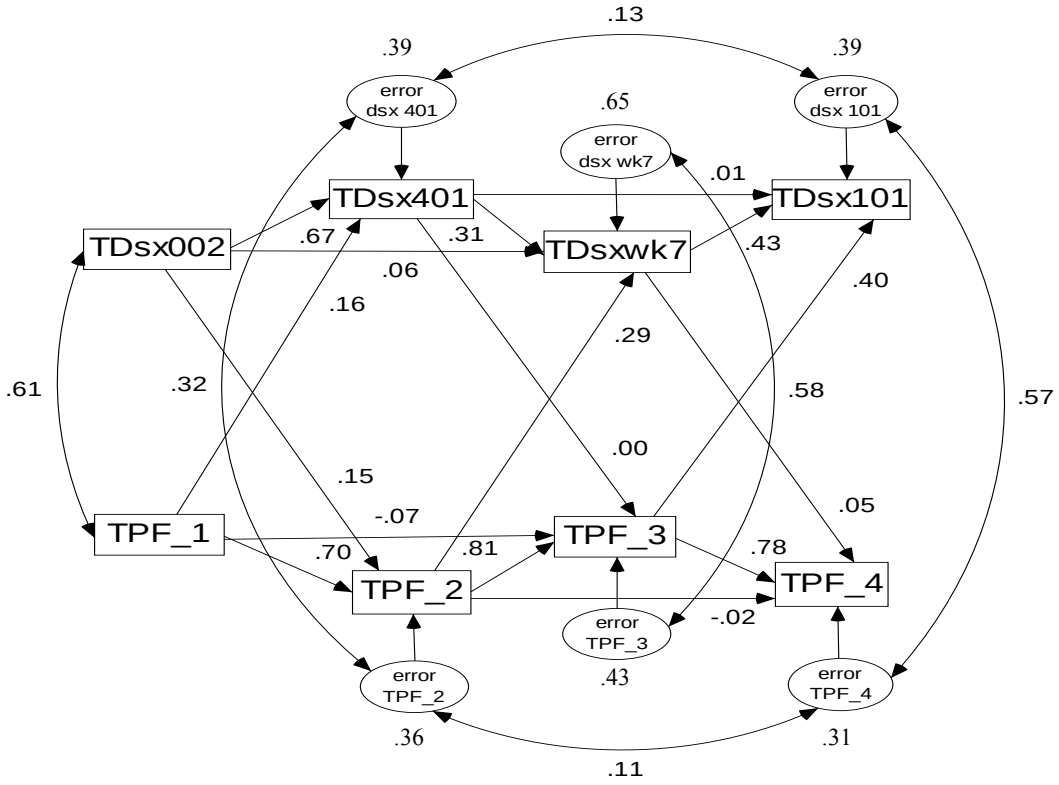


Note. TDsx002 = Index of Depressive Symptoms at Diagnostic Evaluation; TDsx401 = Index of Depressive Symptoms at Week 1 of A-CT; TDsxwk7 = Index of Depressive Symptoms at Week 7 of A-CT; TDsx101 = Index of Depressive Symptoms at Post-A-CT Blind Evaluation; TPF_1 = Index of Psychosocial Functioning at Month of Diagnostic Evaluation; TPF_2 = Index of Psychosocial Functioning at 1st Month of A-CT; TPF_3 = Index of Psychosocial Functioning at 2nd Month of A-CT; TPF_4 = Index of Psychosocial Functioning at 3rd Month of A-CT.

To improve model fit, the paths specifying unexplained sources of variability for the two variables with Heywood cases were trimmed. Specifically, the paths connecting disturbances were rearranged so that only two paths, instead of six, controlled for sources of shared, unexplained variation across repeated measurements. The two remaining paths connected disturbances between the index of depressive symptom severity at week one of A-CT (TDsx401) and the post-A-CT blind evaluation (TDsx101) and the index of psychosocial functioning at month one (TPF_2) and three of A-CT (TPF_4) (see Figure 11). After adjusting these paths, the model no longer had Heywood cases. Yet, model fit still varied, with indices showing good (NFI = 99; CFI = 1.00), moderate (RMSEA = 0.06; 90% CI = 0.04 to 0.10), and poor fit (χ_m^2 [6] = 19.19, p = 0.004).

Figure 11

First Revised Path Analysis Model Showing Mediating Relationships between Psychosocial Functioning and Depressive Symptom Severity: Standardized Solution

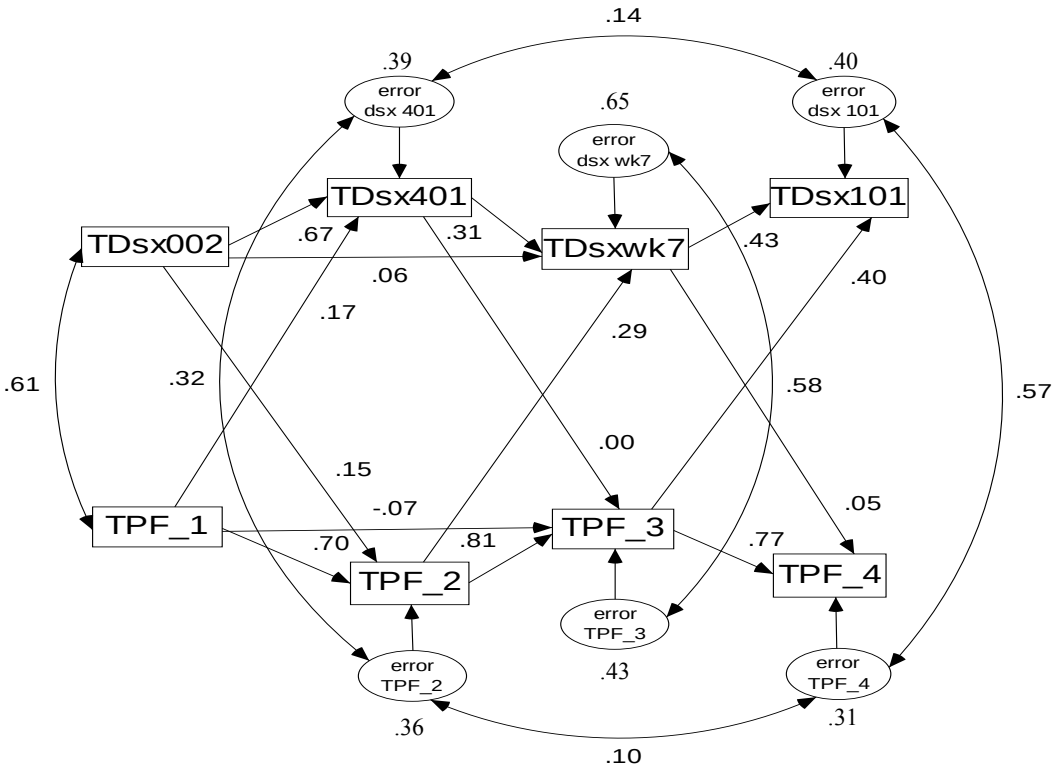


Note. TDsx002 = Index of Depressive Symptoms at Diagnostic Evaluation; TDsx401 = Index of Depressive Symptoms at Week 1 of A-CT; TDsxwk7 = Index of Depressive Symptoms at Week 7 of A-CT; TDsx101 = Index of Depressive Symptoms at Post-A-CT Blind Evaluation; TPF_1 = Index of Psychosocial Functioning at Month of Diagnostic Evaluation; TPF_2 = Index of Psychosocial Functioning at 1st Month of A-CT; TPF_3 = Index of Psychosocial Functioning at 2nd Month of A-CT; TPF_4 = Index of Psychosocial Functioning at 3rd Month of A-CT.

As a next step, non-significant paths were trimmed that did not contribute to model fit. The significance of model paths was determined by comparing critical ratios (obtained by dividing unstandardized direct effects by their standard error) to the critical value of ± 3.2905 for two-tailed statistical significance at the 0.001 level. The contribution of each path deletion to model fit was evaluated with the chi-square difference statistic. If after the path deletion the chi-square difference statistic was not statistically significant ($p < 0.05$), it showed model fit didn't change when the path was trimmed. Two non-significant paths were trimmed between TDsx401 and TDsx101 (χ^2_D [1] = 0.08, $p > 0.05$) and TPF_2 and TPF_4 (χ^2_D [1] = 0.31, $p > 0.05$), as they did not contribute to model fit. However, two non-significant paths (i.e., TDsxwk7 → TPF_4 [χ^2_D (1) = 1.16, $p > 0.05$] and TDsx401 → TPF_3 [χ^2_D (1) = 1.68, $p > 0.05$]) were left in the model, even though they did not contribute to model fit, because they were part of indirect effects tested in the primary aim. All other non-significant paths significantly contributed to model fit. After these two path deletions, only the model chi-square showed poor fit with the data (χ_m^2 [6] = 19.51, p = 0.01; RMSEA = 0.06, 90% CI = 0.02 to 0.09; NFI = 99; CFI = 1.00) (see Figure 12).

Figure 12

Second Revised Path Analysis Model Showing Mediating Relationships between Psychosocial Functioning and Depressive Symptom Severity: Standardized Solution

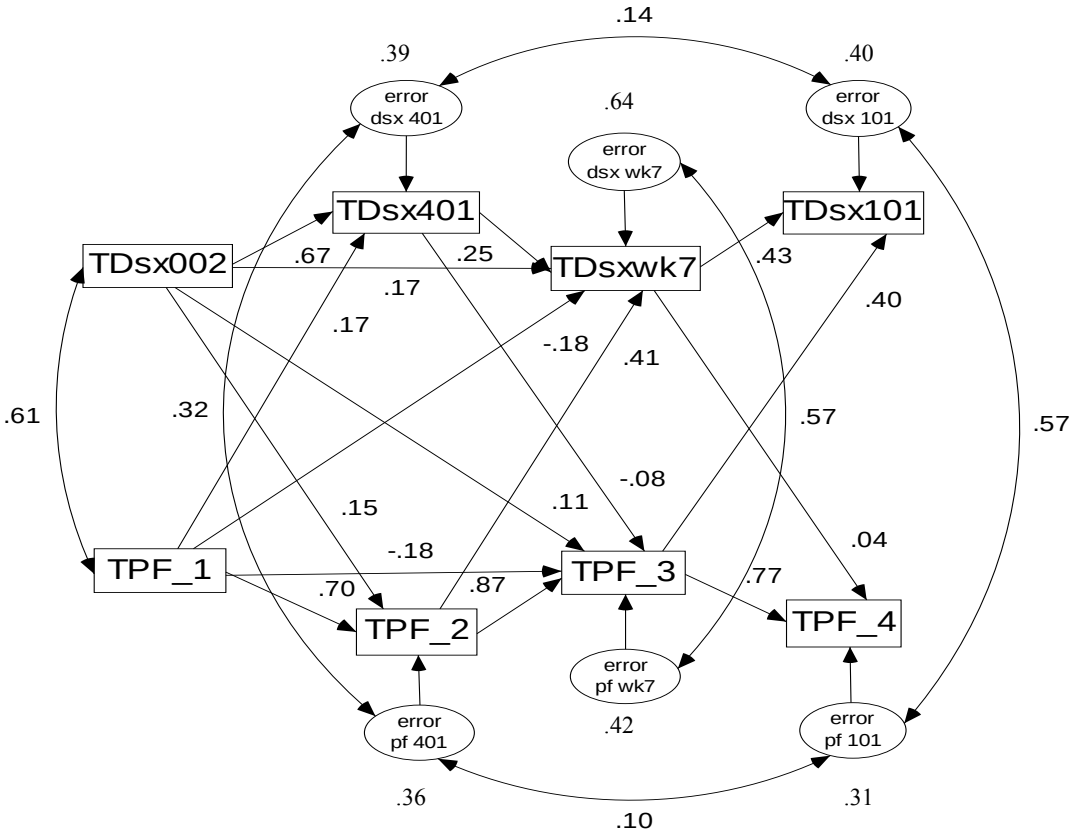


Note. TDsx002 = Index of Depressive Symptoms at Diagnostic Evaluation; TDsx401 = Index of Depressive Symptoms at Week 1 of A-CT; TDsxwk7 = Index of Depressive Symptoms at Week 7 of A-CT; TDsx101 = Index of Depressive Symptoms at Post-A-CT Blind Evaluation; TPF_1 = Index of Psychosocial Functioning at Month of Diagnostic Evaluation; TPF_2 = Index of Psychosocial Functioning at 1st Month of A-CT; TPF_3 = Index of Psychosocial Functioning at 2nd Month of A-CT; TPF_4 = Index of Psychosocial Functioning at 3rd Month of A-CT.

As a final step, paths that lagged across two assessment periods were built to improve model fit. Specifically, paths were added between the index of psychosocial functioning at baseline (TPF_1) and TDsxwk7 (χ^2_D [1] = 4.98, p = 0.05) and the index of depressive symptom severity at baseline (TDsx002) and TPF_3 (χ^2_D [1] = 3.98, p = 0.05) (see Figure 13). The two other potential lag-two paths were not added (i.e., TDsx401 → TPF_4 and TPF_2 → TDsx101) since they did not improve model fit. After adding these two paths, model fit improved to acceptable levels (χ_m^2 [6] = 10.56, p = 0.10; RMSEA = 0.04, 90% CI = 0.00 to 0.08; NFI = 1.00; CFI = 1.00). Since no other path deletion or addition improved model fit, the model in Figure 13 represented the best possible model given theoretical and empirical constraints.

Figure 13

Final Path Analysis Model Showing Mediating Relationships between Psychosocial Functioning and Depressive Symptom Severity: Standardized Solution



Note. TDsx002 = Index of Depressive Symptoms at Diagnostic Evaluation; TDsx401 = Index of Depressive Symptoms at Week 1 of A-CT; TDsxwk7 = Index of Depressive Symptoms at Week 7 of A-CT; TDsx101 = Index of Depressive Symptoms at Post-A-CT Blind Evaluation; TPF_1 = Index of Psychosocial Functioning at Month of Diagnostic Evaluation; TPF_2 = Index of Psychosocial Functioning at 1st Month of A-CT; TPF_3 = Index of Psychosocial Functioning at 2nd Month of A-CT; TPF_4 = Index of Psychosocial Functioning at 3rd Month of A-CT.

Primary Aim: General Evaluation of Parameter Estimates

Before the model in Figure 13 was used to test the primary aim, efforts were first made to understand parameter estimates (i.e., direct effects, error variances, and covariances between indices). As seen in Table 19, after controlling for sources of shared measurement error and inter-correlation between constructs, the model in Figure 13 found 10 significant direct effects (LISREL 8.80 and AMOS 16.0 only determined statistical significance for unstandardized parameter estimates). Unstandardized parameter estimates were interpreted in units of raw data. For example, a one point increase in the index of psychosocial functioning at treatment baseline and month one and two of A-CT led to 0.19, 0.56, and 0.52 point increases in the index of depressive symptom severity at week one and seven of A-CT and the post-A-CT blind evaluation, respectively. So, after controlling for baseline levels of depressive symptom severity, greater impairment in psychosocial functioning led to significantly ($p < 0.001$) worse depressive symptom severity at each lag-one assessment point.

Standardized estimates were interpreted in units of standard deviation. For instance, when the baseline index of depressive symptom severity increased by one standard deviation, the index of psychosocial functioning at month one of A-CT increased by 0.15 standard deviations, after controlling for baseline levels of psychosocial functioning. In other words, worse baseline depressive symptom severity led to greater impairment in psychosocial functioning at the beginning of A-CT.

Table 19

Maximum Likelihood Parameter Estimates from the Final Model for Changes in Psychosocial Functioning and Depressive Symptom Severity: Direct Effects

Parameter	Unstandardized Estimates	Standard Error	Standardized Estimates
TDsx002 → TDsx401	0.764*	0.042	0.669
TPF_1 → TPF_2	0.731*	0.038	0.701
TDsx002 → TPF_2	0.158*	0.038	0.151
TPF_1 → TDsx401	0.190*	0.042	0.166
TDsx002 → TDsxwk7	0.232	0.095	0.165
TDsx401 → TDsxwk7	0.310*	0.085	0.252
TPF_1 → TPF_3	-0.200	0.066	-0.176
TPF_2 → TPF_3	0.951*	0.064	0.874
TDsx002 → TPF_3	0.130	0.063	0.114
TPF_1 → TDsxwk7	-0.256	0.099	-0.182
TDsx401 → TPF_3	-0.077	0.056	-0.077
TPF_2 → TDsxwk7	0.558*	0.097	0.415
TDsxwk7 → TDsx101	0.454*	0.049	0.408
TDsxwk7 → TPF_4	0.038	0.035	0.045
TPF_3 → TPF_4	0.807*	0.047	0.806
TPF_3 → TDsx101	0.520*	0.061	0.424

Note. TDsx002 = Index of Depressive Symptoms at Diagnostic Evaluation; TDsx401 = Index of Depressive Symptoms at Week 1 of A-CT; TDsxwk7 = Index of Depressive Symptoms at Week 7 of A-CT; TDsx101 = Index of Depressive Symptoms at Post-A-CT Blind Evaluation; TPF_1 = Index of Psychosocial Functioning at Month of Diagnostic Evaluation; TPF_2 = Index of Psychosocial Functioning at 1st Month of A-CT; TPF_3 = Index of Psychosocial Functioning at 2nd Month of A-CT; TPF_4 = Index of Psychosocial Functioning at 3rd Month of A-CT.

* $p < 0.001$

The model in Figure 13 also provided data regarding how much variation in each endogenous variable was explained or unexplained. As seen in Table 20, the proportion of unexplained variation for each endogenous variable was estimated in both unstandardized and standardized units. Unstandardized estimates represented unexplained variation in terms of raw data, while standardized estimates signified the proportion of unexplained variance in terms of R^2. For example, the model in Figure 13 did not account for 0.64% of the variation in scores on the index of depressive symptom severity at week seven of A-CT (error_TDsxwk7). On the contrary, the path analysis did account for 36% (i.e., 1- R^2 = 1 – 0.64 = 0.36) of the variability in depressive symptom severity at week seven of A-CT, by accounting for previous levels of depressive symptom severity and psychosocial functioning as well as the covariation between these two constructs. For another example, the path analysis accounted for 69% of the variability in psychosocial functioning during the last month of A-CT (i.e., standardized estimate for unexplained variance [error_TPF_4] = 0.31; 1- R^2 = 1-0.31 = 0.69).

Table 20

Maximum Likelihood Parameter Estimates from the Final Model for Changes in Psychosocial Functioning and Depressive Symptom Severity: Variances and Covariances

Parameter	Unstandardized Estimates	Standard Error	Standardized Estimates
	Variances		
TDsx002	99.787*	6.516	1.00
TPF_1	100.023*	6.641	1.00
error_TDsx401	50.528*	3.374	0.39
error_TPF_2	38.964*	2.664	0.36
error_TDsxwk7	125.738*	9.042	0.64
error_TPF_3	54.561*	3.969	0.42
error_TDsx101	85.539*	6.352	0.40
error_TPF_4	43.041*	3.230	0.31
	Covariances		
TDsx002 ↔ TPF_1	60.840*	5.446	0.609
error_TDsx401 ↔ error_TPF_2	14.342*	2.226	0.323
error_TDsxwk7 ↔ error_TPF_3	47.291*	4.883	0.571
error_TDsx101 ↔ error_TPF_4	34.367*	3.659	0.566
error_TDsx401 ↔ error_TDsx101	9.384	2.877	0.143
error_TPF_2↔ error_TPF_4	4.059	2.020	0.099

Note. TDsx002 = Index of Depressive Symptoms at Diagnostic Evaluation; TDsx401 = Index of Depressive Symptoms at Week 1 of A-CT; TDsxwk7 = Index of Depressive Symptoms at Week 7 of A-CT; TDsx101 = Index of Depressive Symptoms at Post-A-CT Blind Evaluation; TPF_1 = Index of Psychosocial Functioning at Month of Diagnostic Evaluation; TPF_2 = Index of Psychosocial Functioning at 1st Month of A-CT; TPF_3 = Index of Psychosocial Functioning at 2nd Month of A-CT; TPF_4 = Index of Psychosocial Functioning at 3rd Month of A-CT.

* $p < 0.001$

Additionally, the final model in Figure 13 provided data on the degree of covariation between: a) indices of depressive symptom severity and psychosocial functioning at treatment baseline and the beginning, middle, and end of A-CT, and b) sources of shared error variance between indices at the beginning and end of A-CT. While unstandardized estimates represented this covariation in terms of raw data, standardized estimates of this covariation were interpreted like correlation coefficients (*r*). For instance, according to the final model, the correlation coefficients between indices of psychosocial functioning and depressive symptom severity at treatment baseline and the beginning, middle, and end of A-CT were 0.61, 0.32, 0.57, and 0.57, respectively (see Table 20).

Primary Aim: Evaluation of the Partially Mediating Relationships between the Indices of Psychosocial Functioning and Depressive Symptom Severity

Since past research has suggested reductions in depressive symptom severity precede and account for improvements in psychosocial functioning (e.g., Judd, Akiskal, et al., 2000; Hirschfeld et al., 2002; Vittengl et al., 2004), changes in variance unique to the index of depressive symptom severity were predicted to partially mediate changes in variance unique to the index of psychosocial functioning and not vice versa. It was also predicted that the final path analysis model would show significant, independent changes in the indices of psychosocial functioning and depressive symptom severity during A-CT. According to the final model, significant, independent changes in each construct did indeed occur during A-CT. Contrary to what was expected, however, changes in

variance unique to psychosocial functioning partially mediated changes in variance unique to depressive symptom severity and not vice versa.

To reach these conclusions, direct and indirect effects in the final model (see Figure 13) were evaluated and compared. If significant direct effects were noted across repeated measurements for each index, it was determined that significant, independent changes occurred in each construct during A-CT. Additionally, statistically significant indirect effects were compared to the direct effects in the mediating relationship to clarify whether the indirect effect represented a partial or complete mediating effect. If a significant indirect effect was noted in the presence or absence of significant direct effects, then a partial or complete mediating relationship was supported, respectively.

As seen in Table 19, significant direct effects (p = 0.001) linked indices across repeated measurements, showing independent changes occurred in psychosocial functioning and depressive symptom severity during A-CT. Using unstandardized estimates, a one point increase on the index of depressive symptom severity at diagnostic evaluation and week one and seven of A-CT led to 0.76, 0.31, and 0.45 point increases on the index of depressive symptom severity at week one and seven of A-CT and the post-A-CT blind evaluation, respectively. Similarly, unstandardized estimates showed one point increases on the index of psychosocial functioning at diagnostic evaluation and the first and second month of A-CT led to 0.73, 0.95, and 0.81 point increases in the index of psychosocial functioning at month one, two, and three of A-CT, respectively. Taken together, worse depressive symptom severity and psychosocial functioning, evidenced by higher scores on their respective indices, predicted greater impairment at subsequent assessment periods.

The Sobel test was used to test statistical significance of indirect effects between variance unique to psychosocial functioning and depressive symptom severity. In the final model (see Figure 13), three indirect effects tested the degree to which changes in variance unique to depressive symptom severity mediated changes in variance unique to psychosocial functioning (e.g., TPF_1 → TDsx401 → TPF_3; TPF_1 → TDsxwk7 → TPF_4; TPF_2 → TDsxwk7 → TPF_4). Three indirect effects were also available to test the extent to which changes in variance unique to psychosocial functioning mediated changes in variance unique to depressive symptom severity (e.g., TDsx002 → TPF_2 → TDsxwk7, TDsx401 → TPF_3 → TDsx101, and TDsx002 → TPF_3 → TDsx101).

As seen in Table 21, only one indirect effect between variance unique to depressive symptom severity and psychosocial functioning was significant. This indirect effect showed the index of depressive symptom severity at week seven of A-CT increased by 0.09 points for every one point increase in the index of depressive symptom severity at baseline via its prior effect on psychosocial functioning at month one of A-CT. In other words, when controlling for shared sources of measurement error and covariation between constructs, changes in variance unique to psychosocial functioning at month one of A-CT mediated changes in variance unique to depressive symptom severity from baseline to week seven of A-CT.

Table 21

Sobel Test for Statistical Significance of Indirect Effects between Psychosocial Functioning and Depressive Symptom Severity in Final Path Analysis Model

Indirect effect	*a*	SE_a	*b*	SE_b	*ab*	SE_{ab}	*Sobel Statistic*
Psychosocial Functioning as Mediating Variable							
TDsx002 → TPF_2 → TDsxwk7	0.158	0.038	0.558	0.097	0.088	0.026	3.370*
TDsx401 → TPF_3 → TDsx101	-0.077	0.056	0.520	0.061	-0.040	0.031	-1.290
TDsx002 → TPF_3 → TDsx101	0.130	0.063	0.520	0.061	0.068	0.035	2.023
Depressive Symptom Severity as Mediating Variable							
TPF_1 → TDsx401 → TPF_3	0.190	0.042	-0.077	0.056	-0.015	0.011	-1.349
TPF_1 → TDsxwk7 → TPF_4	-0.256	0.099	0.038	0.035	-0.010	0.008	-1.250
TPF_2 → TDsxwk7 → TPF_4	0.558	0.097	0.038	0.035	0.022	0.018	1.222

Note. *a* and SE_a = unstandardized direct effect and standard error for the initial path; *b* and SE_b = unstandardized direct effect and standard error for second path; *ab* = product of unstandardized direct effects or indirect effect; SE_{ab} = standard error for indirect effect (see Sobel, 1986); $z = ab/SE_{ab}$ or Sobel test; TDsx002 = Index of Depressive Symptoms at Diagnostic Evaluation; TDsx401 = Index of Depressive Symptoms at Week 1 of A-CT; TDsxwk7 = Index of Depressive Symptoms at Week 7 of A-CT; TDsx101 = Index of Depressive Symptoms at Post-A-CT Blind Evaluation; TPF_1 = Index of Psychosocial Functioning at Month of Diagnostic Evaluation; TPF_2 = Index of Psychosocial Functioning at 1st Month of A-CT; TPF_3 = Index of Psychosocial Functioning at 2nd Month of A-CT.

* $p < 0.001$

To clarify whether this statistically significant indirect effect represented a partial or complete mediating relationship, it was compared to other direct and indirect effects for the index of depressive symptom severity at week seven of A-CT. As seen in Table 19, the direct effects of baseline depressive symptom severity (p = 0.014) and psychosocial functioning (p = 0.01) on depressive symptom severity at week seven of A-CT did not reach statistical significance at the Bonferroni corrected alpha of 0.001. A second statistically significant indirect effect was found, however, where scores on the index of depressive symptom severity at week one mediated changes in the same index from baseline to week seven of A-CT (a = 0.76; SE_a = 0.04; b = 0.31; SE_b = 0.09; ab = 0.24; SE_{ab} = 0.07; z = 3.58, $p < 0.001$). In other words, the index of depressive symptom severity at week seven of A-CT increased by 0.24 points for every one point increase at baseline via its prior effect on index scores at week one of A-CT. This finding suggested the mediating impact of psychosocial functioning at month one of A-CT on changes in depressive symptom severity from baseline to week seven of A-CT was partial. Furthermore, given the absence of significant baseline direct effects, it appeared that psychosocial functioning and depressive symptom severity at the beginning of A-CT together completely mediated changes in depressive symptom severity from baseline to week seven of A-CT.

Summary of Findings from Primary Aim

A structural equation model was fit with data reflecting psychosocial functioning and depressive symptom severity at treatment baseline and the beginning, middle, and end of A-CT. Given past literature on the topic (e.g., Hirschfeld et al., 2002, Vittengl et

al., 2004), it was hypothesized that change in depressive symptom severity would partially mediate change in psychosocial functioning during A-CT, and not vice versa. Once sources of error variance were controlled (e.g., shared method variance, covariation between constructs), however, it appeared that change in variance unique to psychosocial functioning during the first month of A-CT partially mediated change in variance unique to depressive symptom severity from baseline to week seven of A-CT. Also, after week seven of A-CT, psychosocial functioning predicted subsequent depressive symptom severity, and not vice versa.

Secondary Analyses

Secondary analyses were done to investigate possible reasons why the above finding contradicted the hypothesis and previous research (Finkelstein et al., 1996; Hirschfeld et al., 2002; Lenderking et al., 1999; Vittengl et al., 2004). These secondary analyses suggested that the RIFT, a clinician-rated, longitudinal measure of psychosocial functioning, may have accounted for the difference in findings, as well as this study's use of repeated measurements and structural equation modeling. Details of these analyses follow.

Secondary Analysis 1: Impact of Type of Measurement (Cross-Sectional vs. Longitudinal) and Rater Type (Clinician-Rated vs. Self-Report) on the Mediating Relationship between Psychosocial Functioning and Depressive Symptom Severity

One possible reason why results from the primary analysis differed from previous research was this study's use of the clinician-rated, longitudinal RIFT. Four previous studies investigating change in psychosocial functioning in relation to depressive symptom severity only used self-report measures that assessed functioning cross-sectionally (Finkelstein et al., 1996; Hirschfeld et al., 2002; Lenderking et al., 1999; Vittengl et al., 2004). Interestingly, each of these studies found change in depressive symptom severity partially to completely accounted for change in psychosocial functioning during acute-phase treatment. To better understand the impact of rater (i.e., self-report vs. clinician-rated) and type of measurement (i.e., cross-sectional vs. longitudinal) on the mediating relationship between psychosocial functioning and depressive symptom severity, this study's primary aim was repeated using only the RIFT or SAS-SR, instead of the index of psychosocial functioning. While no mediating effects were found, evidence suggested results of the primary aim were partly explained by the use of the clinician-rated, longitudinal RIFT.

Mediating Relationship between Psychosocial Functioning and Depressive Symptom Severity with RIFT Data

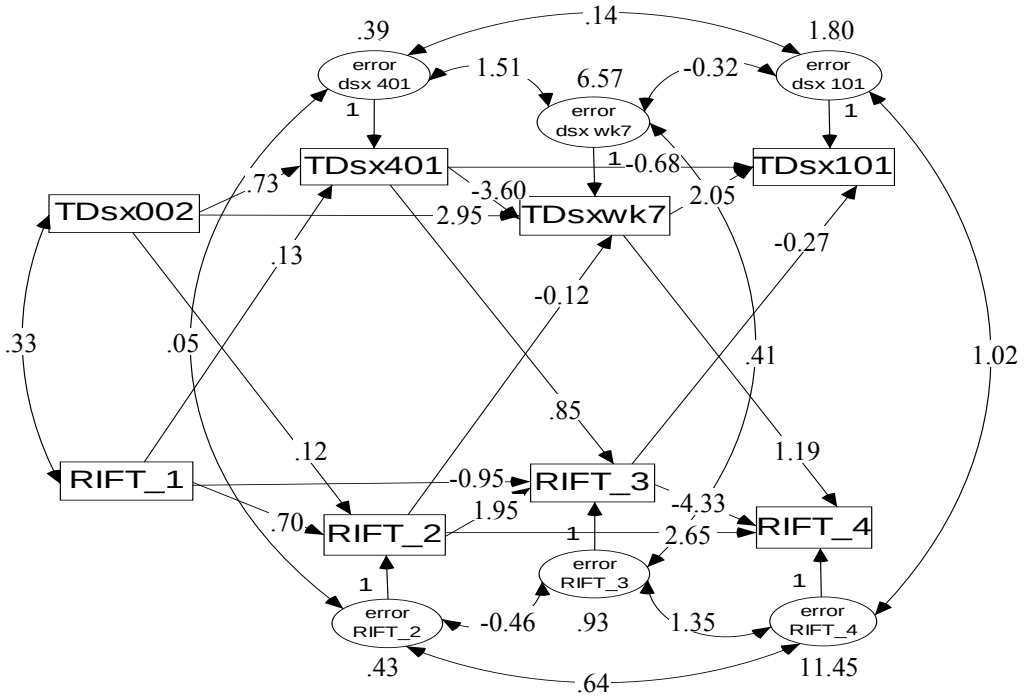
A structural equation model tested the mediating relationship between psychosocial functioning and depressive symptom severity, measured by the RIFT and

index of depressive symptom severity, respectively. As with the primary aim, this structural equation model was fit with data to account for: a) covariation between depressive symptom severity and psychosocial functioning at treatment baseline and the beginning, middle, and end of A-CT, b) shared sources of error variance across repeated measurements, c) change in depressive symptom severity that occurred independently of change in psychosocial functioning, d) change in psychosocial functioning that occurred independently of change in depressive symptom severity, and e) potential mediating relationships between constructs (see Figure 4).

Initially, this model did not converge, meaning the statistical programs (LISREL 8.80 and AMOS 16.0) could not compute parameter estimates after 250 iterative attempts. Hypothetical estimates were provided that showed acceptable fit (χ_m^2 [2] = 2.70, p = 0.26; RMSEA = 0.03, 90% CI = 0.00 to 0.10; CFI = 1.00; NFI = 1.00). As seen in Figure 14, however, Heywood cases were noted for disturbance estimates (cannot exceed 1.0) for the index of depressive symptom severity at week seven of A-CT (6.57) and the post-A-CT blind evaluation (i.e., 1.80), as well as for the RIFT at the end of A-CT (i.e., 11.45). Heywood cases were also noted for covariances (cannot exceed 1.0) between disturbance estimates for: a) the index of depressive symptom severity at week one and seven of A-CT (1.51), b) the RIFT at the second and third month of A-CT (1.35), and c) the index of depressive symptom severity and RIFT at the end of A-CT (1.02).

Figure 14

Initial Model Showing Mediating Relationships between Depressive Symptom Severity and Psychosocial Functioning Measured with the RIFT: Standardized Solution

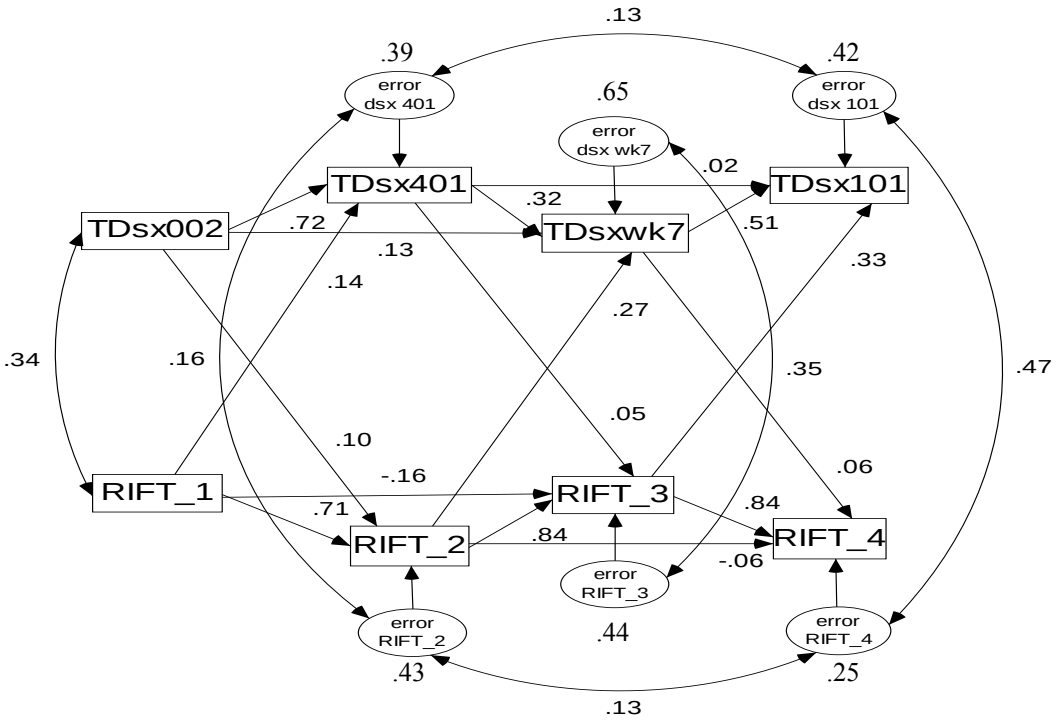


Note. TDsx002 = Index of Depressive Symptoms at Diagnostic Evaluation; TDsx401 = Index of Depressive Symptoms at Week 1 of A-CT; TDsxwk7 = Index of Depressive Symptoms at Week 7 of A-CT; TDsx101 = Index of Depressive Symptoms at Post-A-CT Blind Evaluation; RIFT_1 = Longitudinal Interval Follow-up Evaluation – Range of Impaired Functioning Tool at Month of Diagnostic Evaluation; RIFT_2 = Longitudinal Interval Follow-up Evaluation – Range of Impaired Functioning Tool at 1st Month of A-CT; RIFT_3 = Longitudinal Interval Follow-up Evaluation – Range of Impaired Functioning Tool at 2nd Month of A-CT; RIFT_4 = Longitudinal Interval Follow-up Evaluation – Range of Impaired Functioning Tool at 3rd Month of A-CT.

To improve model fit, paths specifying disturbance estimates (i.e., estimates of unexplained sources of variability) for the six Heywood cases were trimmed so that only two paths, instead of six, controlled for sources of shared, unexplained variation across repeated measurements. The two remaining paths connected disturbances between the index of depressive symptom severity at week one of A-CT (TDsx401) and the post-A-CT blind evaluation (TDsx101) and the RIFT at month one (RIFT_2) and three (RIFT_4) of A-CT (see Figure 15). After trimming these paths, there were no Heywood cases, and model fit was mixed, with two indices showing poor (i.e., χ_m^2 [6] = 35.07, $p < 0.001$; RMSEA = 0.10, 90% CI = 0.07 to 0.14) and good fit (i.e., CFI = 0.98; NFI = 0.99).

Figure 15

Revised Model Showing Mediating Relationships between Depressive Symptom Severity and Psychosocial Functioning Measured with the RIFT: Standardized Solution

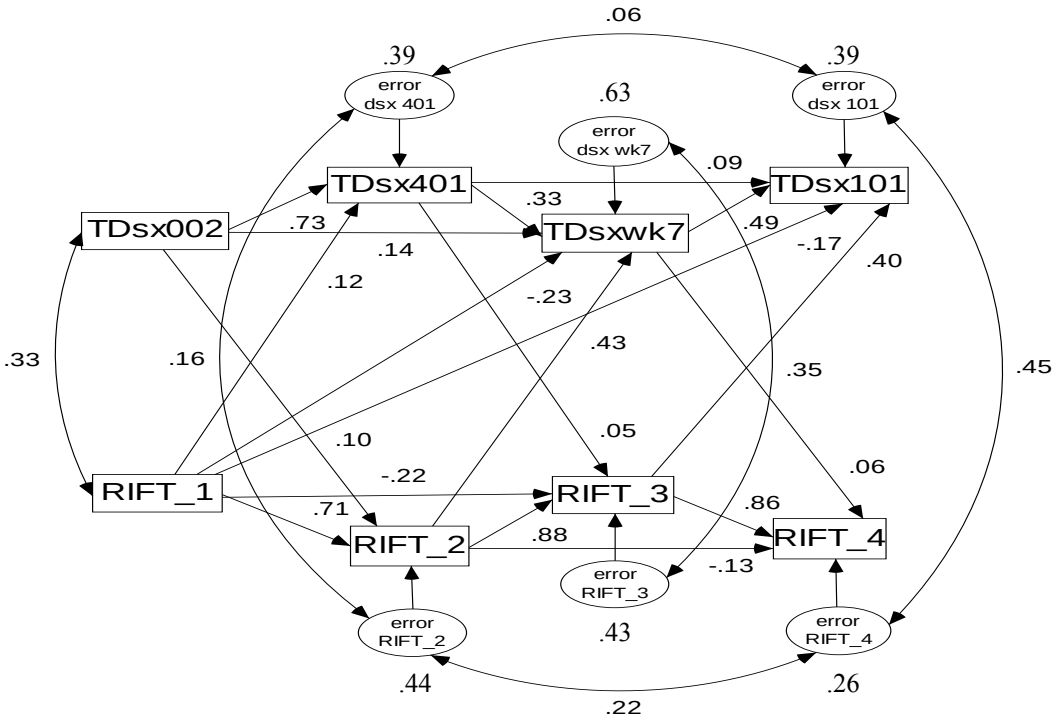


Note. TDsx002 = Index of Depressive Symptoms at Diagnostic Evaluation; TDsx401 = Index of Depressive Symptoms at Week 1 of A-CT; TDsxwk7 = Index of Depressive Symptoms at Week 7 of A-CT; TDsx101 = Index of Depressive Symptoms at Post-A-CT Blind Evaluation; RIFT_1 = Longitudinal Interval Follow-up Evaluation – Range of Impaired Functioning Tool at Month of Diagnostic Evaluation; RIFT_2 = Longitudinal Interval Follow-up Evaluation – Range of Impaired Functioning Tool at 1st Month of A-CT; RIFT_3 = Longitudinal Interval Follow-up Evaluation – Range of Impaired Functioning Tool at 2nd Month of A-CT; RIFT_4 = Longitudinal Interval Follow-up Evaluation – Range of Impaired Functioning Tool at 3rd Month of A-CT.

As a next step, paths were built to improve model fit. Specifically, paths were built between the RIFT at the month of diagnostic evaluations (RIFT_1) and the index of depressive symptom severity at week seven of A-CT (TDsxwk7) (χ^2_D (1) = 11.17, p < 0.001) and RIFT_1 and TDsx101 (χ^2_D (1) = 16.78, p < 0.001) (see Figure 16). After the addition of these paths, model fit improved to acceptable levels (χm^2 [4] = 7.12, p = 0.13; RMSEA = 0.04, 90% CI = 0.00 to 0.09; NFI = 1.00; CFI = 1.00). Of note, two non-significant paths (i.e., TDsx401 → RIFT at month two of A-CT [χ^2_D (1) = 1.86, p > 0.05] and TDsxwk7 → TPF_4 [χ^2_D (1) = 3.31, p > 0.05]) were left in the model, even though they did not contribute to model fit, because they were part of indirect effects tested in this secondary analysis. Since no further path re-specifications significantly improved model fit, the model in Figure 16 represented the best possible model for this analysis, given theoretical and empirical constraints.

Figure 16

Final Model Showing Mediating Relationships between Depressive Symptom Severity and Psychosocial Functioning Measured with the RIFT: Standardized Solution



Note. TDsx002 = Index of Depressive Symptoms at Diagnostic Evaluation; TDsx401 = Index of Depressive Symptoms at Week 1 of A-CT; TDsxwk7 = Index of Depressive Symptoms at Week 7 of A-CT; TDsx101 = Index of Depressive Symptoms at Post-A-CT Blind Evaluation; RIFT_1 = Longitudinal Interval Follow-up Evaluation – Range of Impaired Functioning Tool at Month of Diagnostic Evaluation; RIFT_2 = Longitudinal Interval Follow-up Evaluation – Range of Impaired Functioning Tool at 1st Month of A-CT; RIFT_3 = Longitudinal Interval Follow-up Evaluation – Range of Impaired Functioning Tool at 2nd Month of A-CT; RIFT_4 = Longitudinal Interval Follow-up Evaluation – Range of Impaired Functioning Tool at 3rd Month of A-CT.

After the model showed acceptable fit with the data, significant direct and indirect effects were examined to determine the mediating relationship between psychosocial functioning, measured by the RIFT, and depressive symptom severity. As seen in Table 22, when sources of shared measurement error and inter-correlation between constructs were controlled, the model in Figure 16 showed 11 significant direct effects. According to unstandardized parameter estimates, a one point increase on the RIFT at treatment baseline and month one and two of A-CT led to 0.51, 2.07, and 1.84 point increases in the index of depressive symptom severity at week one and seven of A-CT and the post-A-CT blind evaluation, respectively. So, after controlling for baseline levels of depressive symptom severity, greater impairment in psychosocial functioning led to significantly worse depressive symptom severity at each lag-one assessment point. Change in the depressive symptom severity did not significantly predict change in psychosocial functioning, and there were no significant indirect effects (see Table 23).

Table 22

Maximum Likelihood Direct Effects from Final Model for Changes in Depressive Symptom Severity and Psychosocial Functioning Measured by the RIFT

Parameter	Unstandardized Estimates	Standard Error	Standardized Estimates
TDsx002 → TDsx401	0.833*	0.036	0.730
RIFT_1 → RIFT_2	0.742*	0.039	0.712
TDsx002 → RIFT_2	0.028	0.011	0.099
RIFT_1 → TDsx401	0.506*	0.146	0.122
TDsx002 → TDsxwk7	0.199	0.083	0.144
TDsx401 → TDsxwk7	0.399*	0.077	0.328
RIFT_1 → RIFT_3	-0.245*	0.059	-0.220
RIFT_2 → RIFT_3	0.938*	0.058	0.880
RIFT_1 → TDsxwk7	-1.137	0.321	-0.227
TDsx401 → RIFT_3	0.015	0.010	0.055
RIFT_2 → TDsxwk7	2.072*	0.310	0.430
TDsxwk7 → TDsx101	0.501*	0.045	0.489
TDsxwk7 → RIFT_4	0.015	0.008	0.064
RIFT_3 → RIFT_4	0.916*	0.047	0.860
RIFT_3 → TDsx101	1.841*	0.198	0.398
TDsx401→ TDsx101	0.107	0.061	0.086
RIFT_2 → RIFT_4	-0.149	0.053	-0.131
RIFT_1 → TDsx101	-0.853*	0.203	-0.166

Note. TDsx002 = Index of Depressive Symptoms at Diagnostic Evaluation; TDsx401 = Index of Depressive Symptoms at Week 1 of A-CT; TDsxwk7 = Index of Depressive Symptoms at Week 7 of A-CT; TDsx101 = Index of Depressive Symptoms at Post-A-CT Blind Evaluation; RIFT_1 = Longitudinal Interval Follow-up Evaluation – Range of Impaired Functioning Tool at Month of Diagnostic Evaluation; RIFT_2 = Longitudinal Interval Follow-up Evaluation – Range of Impaired Functioning Tool at 1st Month of A-CT; RIFT_3 = Longitudinal Interval Follow-up Evaluation – Range of Impaired Functioning Tool at 2nd Month of A-CT; RIFT_4 = Longitudinal Interval Follow-up Evaluation – Range of Impaired Functioning Tool at 3rd Month of A-CT.

* $p < 0.001$

Table 23

Sobel Test for Statistical Significance of Indirect Effects between Depressive Symptom Severity and Psychosocial Functioning Measured by the RIFT

Indirect effect	*a*	SE_a	*b*	SE_b	*ab*	SE_{ab}	*Sobel Statistic*
Psychosocial Functioning as Mediating Variable							
TDsx002 → RIFT_2 → TDsxwk7	0.028	0.011	2.072	0.31	0.058	0.024	2.379
TDsx401 → RIFT_3 → TDsx101	0.015	0.01	1.841	0.198	0.028	0.019	1.481
Depressive Symptom Severity as Mediating Variable							
RIFT_1 → TDsx401 → RIFT_3	0.506	0.146	0.015	0.010	0.008	0.006	1.377
RIFT_1 → TDsxwk7 → RIFT_4	-1.137	0.321	0.015	0.008	-0.017	0.010	-1.657
RIFT_2 → TDsxwk7 → RIFT_4	2.072	0.310	0.015	0.008	0.031	0.017	1.805

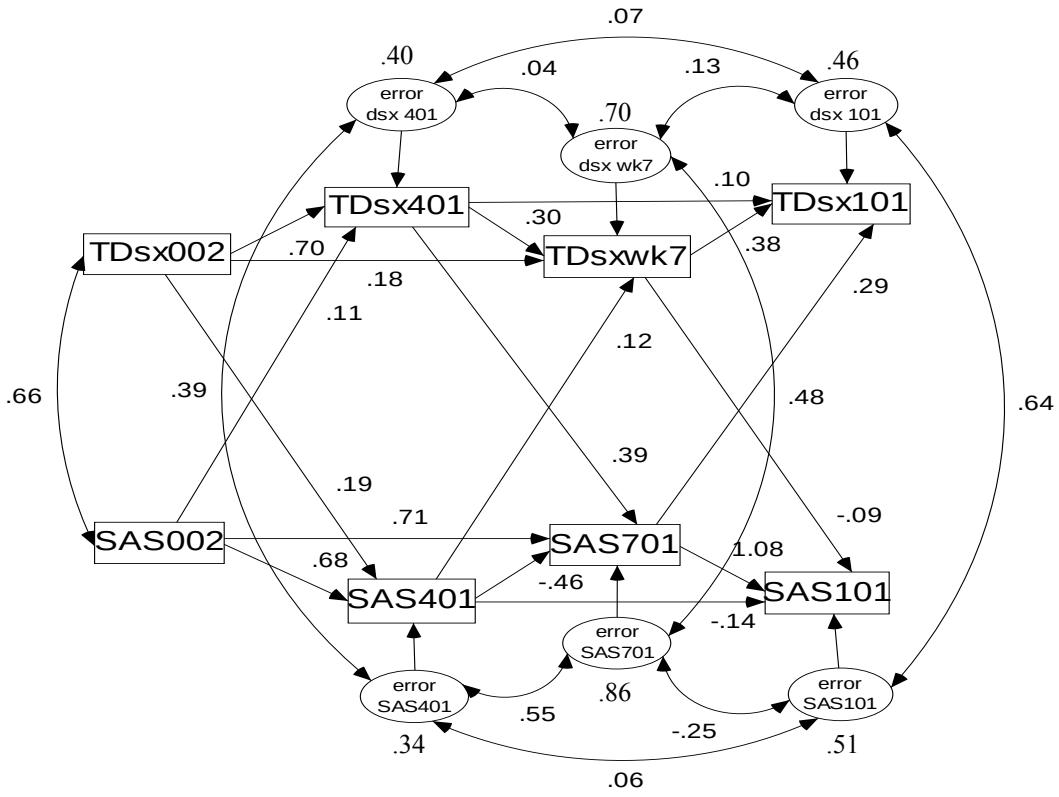
Note. *a* and SE_a = unstandardized direct effect and standard error for the initial path; *b* and SE_b = unstandardized direct effect and standard error for second path; *ab* = product of unstandardized direct effects or indirect effect; SE_{ab} = standard error for indirect effect (see Sobel, 1986); $z = ab/SE_{ab}$ or Sobel test; TDsx002 = Index of Depressive Symptoms at diagnostic evaluation; TDsx401 = Index of Depressive Symptoms at week 1 of A-CT; TDsxwk7 = Index of Depressive Symptoms at week 7 of A-CT; TDsx101 = Index of Depressive Symptoms at post-A-CT blind evaluation; RIFT_1 = Longitudinal Interval Follow-up Evaluation – Range of Impaired Functioning Tool at month of diagnostic evaluations; RIFT_2 = Longitudinal Interval Follow-up Evaluation – Range of Impaired Functioning Tool at 1st month of A-CT; RIFT_3 = Longitudinal Interval Follow-up Evaluation – Range of Impaired Functioning Tool at 2nd month of A-CT; RIFT_4 = Longitudinal Interval Follow-up Evaluation – Range of Impaired Functioning Tool at 3rd month of A-CT.

Mediating Relationship between Depressive Symptom Severity and Psychosocial Functioning Measured with the SAS-SR

The above analyses were repeated to examine the mediating relationship between psychosocial functioning, measured by the self-report, cross-sectional SAS-SR, and depressive symptom severity, measured by the index of depressive symptoms severity. As before, the structural equation model was fit with data to account for: a) covariation between depressive symptom severity and psychosocial functioning at treatment baseline and the beginning, middle, and end of A-CT, b) shared sources of error variance across repeated measurements, c) change in depressive symptom severity that occurred independently of change in psychosocial functioning, d) change in psychosocial functioning that occurred independently of change in depressive symptom severity, and e) potential mediating relationships between constructs (see Figure 4). As seen in Figure 17, the structural equation model initially showed acceptable fit with the data (χ_m^2 [2] = 1.15, p = 0.56; RMSEA = 0.00, 90% CI = 0.00 to 0.08; CFI = 1.00; NFI = 1.00).

Figure 17

Initial Model Showing Mediating Relationships between Depressive Symptom Severity and Psychosocial Functioning Measured with the SAS-SR: Standardized Solution

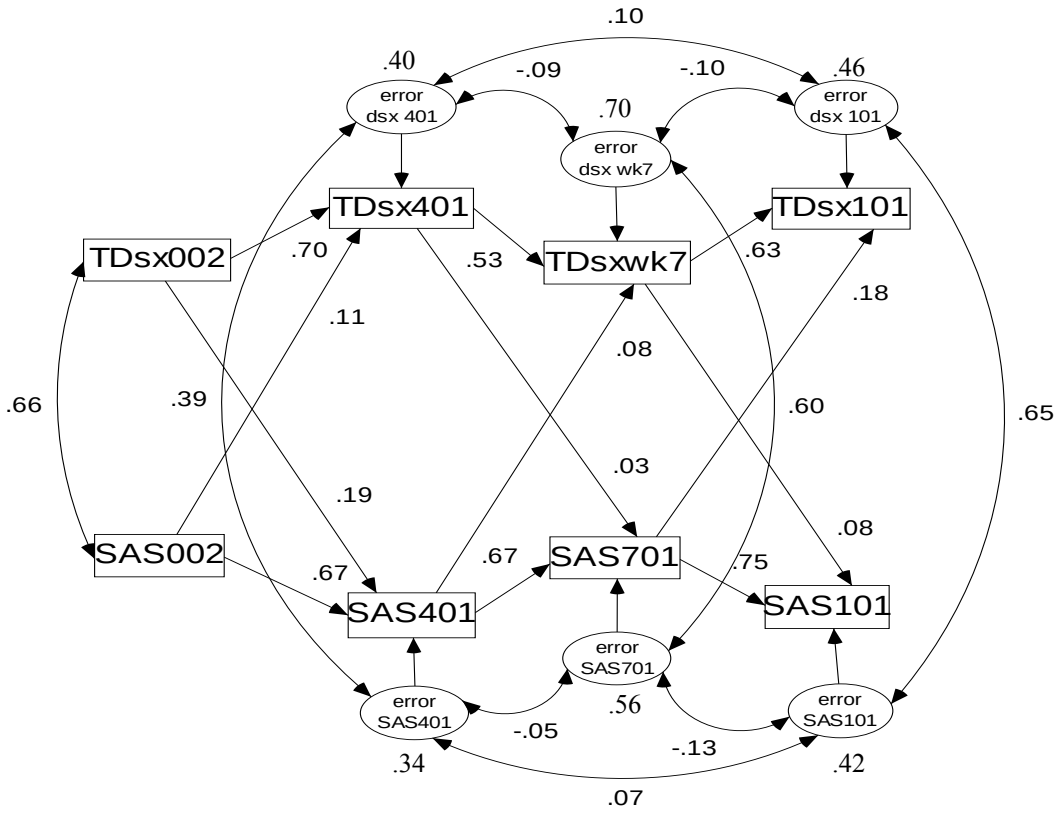


Note. TDsx002 = Index of Depressive Symptoms at Diagnostic Evaluation; TDsx401 = Index of Depressive Symptoms at Week 1 of A-CT; TDsxwk7 = Index of Depressive Symptoms at Week 7 of A-CT; TDsx101 = Index of Depressive Symptoms at Post-A-CT Blind Evaluation; SAS002 = Social Adjustment Scale-Self Report at Diagnostic Evaluation; SAS401 = Social Adjustment Scale-Self Report at week one of A-CT; SAS701 = Social Adjustment Scale-Self Report at week seven of A-CT; SAS101 = Social Adjustment Scale-Self Report at post-A-CT blind evaluation.

To obtain the best model possible (i.e., model that is parsimonious and still fits the data), non-significant paths were eliminated that did not contributed to model fit. Specifically, four non-significant paths were trimmed (i.e., TDsx002 → TDsxwk7 [χ^2_D [1] = 0.46, $p > 0.05$], TDsx401 → TDsx101 [χ^2_D [1] = 2.00, $p > 0.05$], SAS-SR at diagnostic evaluation → SAS-SR at week seven of A-CT [χ^2_D [1] = 0.57, $p > 0.05$], SAS-SR at week one of A-CT → SAS-SR at post-A-CT blind evaluation [χ^2_D [1] = 2.22, $p > 0.05$]), as they did not contribute to model fit. Three non-significant paths that did not contribute to model fit (i.e., TDsx401 → SAS-SR at week seven of A-CT [χ^2_D (1) = 0.20, $p > 0.05$], SAS-SR at week one of A-CT → TDsxwk7 [χ^2_D (1) = 1.68, $p > 0.05$], TDsxwk7 → SAS-SR at post-A-CT blind evaluation [χ^2_D (1) = 1.44, $p > 0.05$]) were left in the model because they were part of indirect effects tested in this secondary analysis. All other non-significant paths significantly contributed to model fit. After these four path deletions, the model again showed acceptable fit with the data (χm^2 [6] = 6.39, p = 0.38; RMSEA = 0.01, 90% CI = 0.00 to 0.06; NFI = 1.00; CFI = 1.00) (see Figure 18). Since no other path deletion or addition significantly improved model fit, the model in Figure 18 represented the best possible model for this analysis, given theoretical and empirical constraints.

Figure 18

Final Model Showing Mediating Relationships between Depressive Symptom Severity and Psychosocial Functioning Measured with the SAS-SR: Standardized Solution



Note. TDsx002 = Index of Depressive Symptoms at Diagnostic Evaluation; TDsx401 = Index of Depressive Symptoms at Week 1 of A-CT; TDsxwk7 = Index of Depressive Symptoms at Week 7 of A-CT; TDsx101 = Index of Depressive Symptoms at Post-A-CT Blind Evaluation; SAS002 = Social Adjustment Scale-Self Report at Diagnostic Evaluation; SAS401 = Social Adjustment Scale-Self Report at week one of A-CT; SAS701 = Social Adjustment Scale-Self Report at week seven of A-CT; SAS101 = Social Adjustment Scale-Self Report at post-A-CT blind evaluation.

As seen in Table 24, after controlling for sources of shared measurement error and inter-correlation between constructs, the model in Figure 18 showed seven significant direct effects. Of these, six direct effects predicted independent change within each construct, and only one direct effect predicted change across constructs, such that a one point increase on the index of depressive symptom severity at diagnostic evaluation predicted a 0.01 point increase on the SAS-SR at the first week of A-CT. In other words, greater depressive symptom severity at diagnostic evaluation predicted slightly worse psychosocial functioning at the beginning of A-CT. When measured by the SAS-SR, change in psychosocial functioning did not predict subsequent depressive symptom severity, and there were no significant indirect effects between constructs (see Table 25).

Table 24

Maximum Likelihood Direct Effects from Final Model for Changes in Depressive Symptom Severity and Psychosocial Functioning Measured by the SAS-SR

Parameter	Unstandardized Estimates	Standard Error	Standardized Estimates
TDsx002 → TDsx401	0.799*	0.045	0.701
SAS002 → SAS401	0.681*	0.039	0.192
TDsx002 → SAS401	0.009*	0.002	0.192
SAS002 → TDsx401	2.734	1.032	0.107
TDsx401 → TDsxwk7	0.647*	0.101	0.529
SAS401 → SAS701	0.694*	0.073	0.668
TDsx401 → SAS701	0.001	0.003	0.027
SAS401 → TDsxwk7	2.542	2.175	0.082
TDsxwk7 → TDsx101	0.656*	0.134	0.631
TDsxwk7 → SAS101	0.003	0.002	0.085
SAS701 → SAS101	0.718*	0.087	0.745
SAS701 → TDsx101	5.545	2.926	0.180

Note. TDsx002 = Index of Depressive Symptoms at Diagnostic Evaluation; TDsx401 = Index of Depressive Symptoms at Week 1 of A-CT; TDsxwk7 = Index of Depressive Symptoms at Week 7 of A-CT; TDsx101 = Index of Depressive Symptoms at Post-A-CT Blind Evaluation; SAS002 = Social Adjustment Scale-Self Report at Diagnostic Evaluation; SAS401 = Social Adjustment Scale-Self Report at week one of A-CT; SAS701 = Social Adjustment Scale-Self Report at week seven of A-CT; SAS101 = Social Adjustment Scale-Self Report at post-A-CT blind evaluation.

* $p < 0.001$

Table 25

Sobel Test for Statistical Significance of Indirect Effects between Depressive Symptom Severity and Psychosocial Functioning Measured by the SAS-SR

Indirect effect	*a*	SE_a	*b*	SE_b	*ab*	SE_{ab}	*Sobel Statistic*
Psychosocial Functioning as Mediating Variable							
TDsx002 → SAS401 → TDsxwk7	0.009	0.002	2.542	2.175	0.023	0.020	1.131
TDsx401 → SAS701 → TDsx101	0.001	0.003	5.545	2.926	0.006	0.017	0.328
Depressive Symptom Severity as Mediating Variable							
SAS002 → TDsx401 → SAS701	2.734	1.032	0.001	0.003	0.003	0.008	0.331
SAS401 → TDsxwk7 → SAS101	2.542	2.175	0.003	0.002	0.008	0.008	0.922

Note. *a* and SE_a = unstandardized direct effect and standard error for the initial path; *b* and SE_b = unstandardized direct effect and standard error for second path; *ab* = product of unstandardized direct effects or indirect effect; SE_{ab} = standard error for indirect effect (see Sobel, 1986); $z = ab/SE_{ab}$ or Sobel test; TDsx002 = Index of Depressive Symptoms at Diagnostic Evaluation; TDsx401 = Index of Depressive Symptoms at Week 1 of A-CT; TDsxwk7 = Index of Depressive Symptoms at Week 7 of A-CT; TDsx101 = Index of Depressive Symptoms at Post-A-CT Blind Evaluation; SAS002 = Social Adjustment Scale-Self Report at Diagnostic Evaluation; SAS401 = Social Adjustment Scale-Self Report at week one of A-CT; SAS701 = Social Adjustment Scale-Self Report at week seven of A-CT; SAS101 = Social Adjustment Scale-Self Report at post-A-CT blind evaluation.

Summary of Secondary Analysis 1

According to results, no significant mediating relationships were detected between psychosocial functioning and depressive symptom severity once either the RIFT or SAS-SR was used to operationalize psychosocial functioning. Nevertheless, greater impairment in psychosocial functioning, measured by the clinician-rated, longitudinal RIFT, predicted significantly higher depressive symptom severity. In contrast, the constructs had little, if any, impact on the other, when psychosocial functioning was measured by the self-report, cross-sectional SAS-SR. Therefore, it could be concluded that results from the primary aim were, at least in part, explained by the use of clinician-rated and/or longitudinal data.

Secondary Analysis 2: Impact of Number of Assessment Points on the Mediating Relationship Between Psychosocial Functioning and Depressive Symptom Severity

A second reason why results from this study's primary aim may have differed from previous literature was its inclusion of change data from four assessment points (i.e., treatment baseline and the beginning, middle, and end of A-CT). By way of comparison, past studies of change in psychosocial functioning and depressive symptom severity only examined change between two assessment points (i.e., pre- and post-treatment; Finkelstein et al., 1996; Hirschfeld et al., 2002; Lenderking et al., 1999; Vittengl et al., 2004). As a result, these past studies could have been less sensitive to subtle changes in psychosocial functioning and depressive symptom severity that occurred between shorter intervals during the acute-phase. To test this claim, the

structural equation model from the primary aim was repeated, using data from only two assessment points (i.e., pre- to post-treatment). From results, it could be concluded that the mediating relationship between psychosocial functioning and depressive symptom severity was influenced by the inclusion of data from more than two assessment points.

Mediating Relationship between Psychosocial Functioning and Depressive Symptom Severity with Only Pre- to Post-Treatment Data

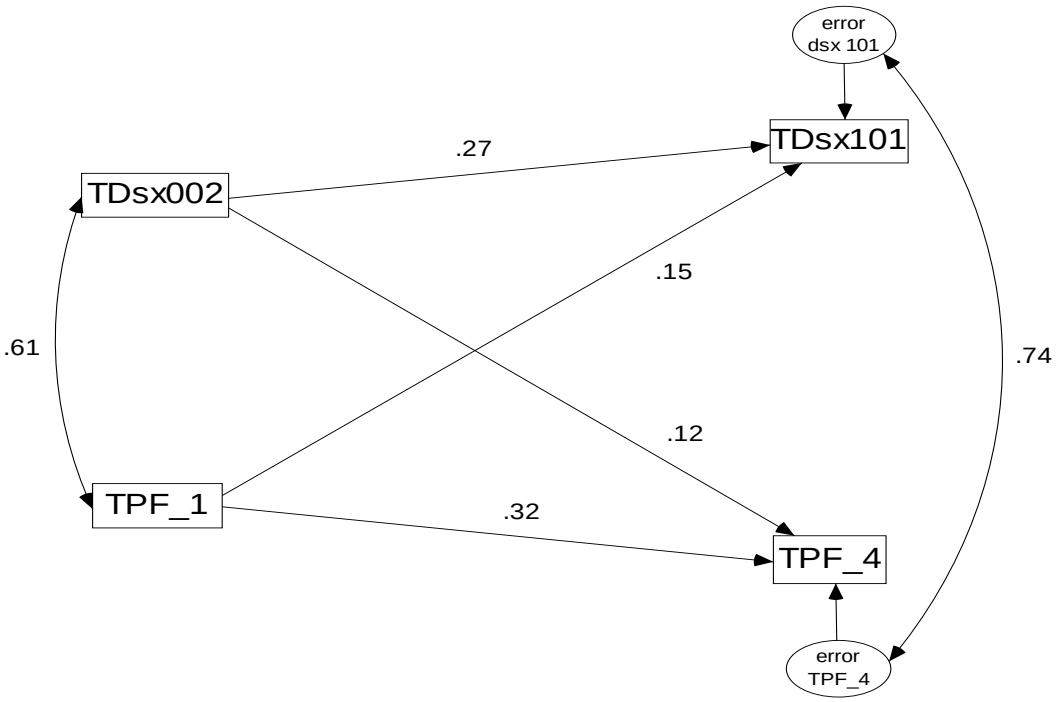
The structural equation model from the primary aim (see Figure 4) was repeated, using only pre- and post-A-CT data. In other words, data from the beginning and middle of A-CT were omitted, leaving only psychosocial functioning and depressive symptom severity data from the diagnostic evaluation and end of A-CT. Constructs were operationalized by the indices of psychosocial functioning and depressive symptom severity.

Since the proposed structural equation model only had four observed variables (i.e., v), the number of parameters that could be estimated equaled 10 (i.e., maximum number of parameters equals number of observations, which equals $v[v + 1]/2 = 4\ [5]\ /2 = 10$, see p. 116 for more details). As a result, the model was fit with data to account for: a) covariation between depressive symptom severity and psychosocial functioning at treatment baseline and the end of A-CT, b) change in depressive symptom severity that occurred independently of change in psychosocial functioning, c) change in psychosocial functioning that occurred independently of change in depressive symptom severity, and d) predictive relationships between constructs. As seen in Figure 19, the model was just-identified (i.e., number of parameter estimates equaled number of observations), so fit

was optimal (χ_m^2 = 0.00, p = 1.00; RMSEA = 0.00, 90% CI = 0.00 to 0.00; CFI = 1.00; NFI = 1.00). No path deletions were made to the model in Figure 19, since data relevant to this secondary analysis would have been lost.

Figure 19

Just-Identified Model Showing Relationship between Psychosocial Functioning and Depressive Symptom Severity: Standardized Solution



Note. TDsx002 = Index of Depressive Symptoms at Diagnostic Evaluation; TDsx101 = Index of Depressive Symptoms at Post-A-CT Blind Evaluation; TPF_1 = Index of Psychosocial Functioning at Month of Diagnostic Evaluation; TPF_4 = Index of Psychosocial Functioning at 3rd Month of A-CT.

As seen in Table 26, after controlling for the inter-correlation between constructs, significant, independent changes occurred in each construct. In other words, a one point increase in baseline indices of psychosocial functioning and depressive symptom severity predicted 0.36 and 0.38 point increases in each respective index after A-CT. Thus, more

impairment in baseline psychosocial functioning and depressive symptom severity predicted worse impairment in each construct after A-CT.

When only pre- to post-A-CT data was examined, however, baseline levels of psychosocial functioning did not significantly predict levels of depressive symptom severity at the end of A-CT, or vice versa. Consequently, increasing the number of assessment points during acute-phase treatment from two to four may have, in part, accounted for the difference between this and previous studies' results. In other words, by splitting the acute-phase up into smaller intervals, this study may have been better able to detect mediating effects between psychosocial functioning and depressive symptom severity.

Table 26

Maximum Likelihood Direct Effects from Model of Pre- to Post-Acute-Phase Cognitive Therapy Changes in Psychosocial Functioning and Depressive Symptom Severity

Parameter	Unstandardized Estimates	Standard Error	Standardized Estimates
TDsx002 → TDsx101	0.379*	0.086	0.272
TPF_1 → TPF_4	0.358*	0.068	0.321
TDsx002 → TPF_4	0.139	0.068	0.124
TPF_1 → TDsx101	0.209	0.086	0.151

Note. TDsx002 = Index of Depressive Symptoms at Diagnostic Evaluation; TDsx101 = Index of Depressive Symptoms at Post-A-CT Blind Evaluation; TPF_1 = Index of Psychosocial Functioning at Month of Diagnostic Evaluation; TPF_4 = Index of Psychosocial Functioning at 3rd Month of A-CT.

* $p < 0.001$

Secondary Analysis 3: Structural Equation Modeling vs. Linear Regression

Finally, results from this study may have differed from previous literature because this study used structural equation modeling, while previous analyses used linear regression (e.g., Finkelstein et al., 1996; Hirschfeld et al., 2002; Lenderking et al., 1999; Vittengl et al., 2004). Regarding this issue, Kline (2005) reported that structural equation modeling was favored over linear regression in mediational analyses because it: a) allowed variables to be entered as both predictors and criterion in the same analysis, and b) controlled for inter-correlation between variables. To test the hypothesis that differences in statistical approaches accounted for this and past studies' divergent results, linear regressions were run to examine the extent to which change in psychosocial functioning accounted for change in depressive symptom severity and vice versa. Constructs were operationalized using the indices of psychosocial functioning and depressive symptom severity, and change was calculated by subtracting pre-A-CT index scores from post-A-CT index scores. The *t*–test for the non-zero intercept represented change in the criterion variable that was independent of the predictor. Results showed that the difference in statistical approaches may have accounted for this and past studies' divergent findings.

Investigating the Relationship between Psychosocial Functioning and Depressive Symptom Severity with Linear Regression

As seen in Tables 27 and 28, linear regression analyses showed change in psychosocial functioning and depressive symptom severity predicted significant change

in the other construct ($t = 14.83$, $p < 0.001$). Based on the non-zero intercept *t*–test, while depressive symptom severity changed independently of changes in psychosocial functioning during A-CT ($t = 16.17$, $p < 0.001$; see Table 28), psychosocial functioning did not change independently of changes in depressive symptom severity ($t = 1.42$, $p = 0.16$; see Table 27). In other words, according to linear regression analyses, change in depressive symptom severity completely accounted for change in psychosocial functioning, while change in psychosocial functioning only partially accounted for change in depressive symptom severity. When this finding was compared to results from the primary aim (i.e., change in psychosocial functioning mediated change in depressive symptom severity and not vice versa), evidence suggested that divergent statistical analyses may have indeed attributed, at least in part, to the difference in findings from this and past studies.

Table 27

Linear Regression Predicting Changes in Psychosocial Functioning from Changes in Depressive Symptom Severity

Model		Unstandardized Coefficients		Standardized Coefficients		
		B	Std. Error	Beta	t	Sig.
1	(Constant)	1.504	1.063		1.42	0.16
	Change in Dep. Sx.	0.516	0.035	0.620	14.83	0.000

Note. Dep. Sx. = Index of Depressive Symptom Severity.

Table 28

Linear Regression Predicting Changes in Depressive Symptom Severity from Changes in Psychosocial Functioning

Model		Unstandardized Coefficients		Standardized Coefficients		
		B	Std. Error	Beta	t	Sig.
1	(Constant)	15.677	0.970		16.17	0.000
	Change in PF	0.744	0.050	0.620	14.83	0.000

Note. PF = Index of Psychosocial Functioning.

CHAPTER SIX
Discussion

This dissertation uses structural equation modeling to test a theoretical framework conceptualizing when and how changes occur in psychosocial functioning during A-CT relative to changes in depressive symptom severity. As predicted, psychosocial functioning and depressive symptom severity show significant, independent changes throughout A-CT. Contrary to what was hypothesized, however, change in depressive symptom severity does not mediate change in psychosocial functioning. Instead, change in psychosocial functioning during the first month of A-CT partially mediates change in depressive symptom severity that occurs between treatment baseline and week seven of A-CT. After week seven of A-CT, greater impairment in psychosocial functioning predicts higher subsequent depressive symptom severity.

In terms of the theoretical framework developed for this study, when patients are exposed to the environmental influence of A-CT, significant, independent improvements occur in their behavior (e.g., psychosocial functioning) and personal factors (e.g., depressive symptom severity). Furthermore, during the first seven weeks of A-CT, changes in behavior partially mediate changes in personal factors and not vice versa. These results add knowledge to the field, by disentangling to some degree the process by which psychosocial functioning and depressive symptom severity change during A-CT in individuals suffering from MDD. To better understand this dissertation's potential contribution to the field, this discussion takes a closer look at the context of these results and compares them to existing literature. Findings are discussed in the order they are

presented above, beginning with an exploration of the psychometric quality of the indices of psychosocial functioning and depressive symptom severity.

Psychometric Quality of the Indices of Psychosocial Functioning and Depressive Symptom Severity

Research showed measures of psychosocial functioning assessed the construct differently and varied in their degree of overlap with measures of depressive symptom severity (e.g., Dunn et al., 2007; Vittengl et al., 2004; Weissman et al., 2001). Therefore, indices of these two constructs were created for the primary aim, combining data from multiple measures and informant types. As a result, this study expected to: a) better isolate treatment effects by accounting for sources of measurement error (e.g., shared method variance) and b) improve the reliability and validity with which constructs were assessed (Kline, 2005). Past research supported the internal consistency reliability of the index of depressive symptom severity, and its utility in measuring pre- to post-A-CT changes (Vittengl et al., 2004). However, the index of psychosocial functioning had not been used previously in the literature. Therefore, this study sought to psychometrically validate the index of psychosocial functioning and replicate previous findings for the index of depressive symptom severity. Results found each index was sufficiently reliable and valid in assessing their respective constructs to be included in the primary aim.

Index Internal Consistency Reliability

Overall, the indices of psychosocial functioning (median $\alpha = 0.65$) and depressive symptom severity (median $\alpha = 0.88$) demonstrated acceptable levels of

internal consistency reliability. At the beginning of the acute-phase, however, reliability coefficients for the index of psychosocial functioning were low (α at month of diagnostic evaluations = 0.49; α at first month of A-CT = 0.57). To interpret this finding, three issues were explored.

First, reliability coefficients for initial scores on the index of psychosocial functioning could have been low due to state-dependent memory biases on the RIFT (see appendix G for psychometric quality of RIFT). Cognitive researchers have reported that "when the mood of the events or material during the learning episode is matched by the current mood, the result is better memory performance during memory retrieval" (p. 394; Barry, Naus, & Rehm, 2004). This was most often not the case in the current study, as most patients reported having a different mood at the end of A-CT, when they completed the RIFT, compared to the beginning of A-CT. For example, some patients were depressed at treatment baseline but not when the RIFT was administered. Therefore, patients may have had difficulty retrieving events from when their mood was more depressed, potentially lowering the RIFT's reliability in assessing psychosocial functioning during the month of diagnostic evaluations and the first month of A-CT (Cole et al., 2003). In other words, mismatched mood states at baseline and RIFT administration could contribute to lowered reliability.

Second, Cronbach's alpha has been associated with the number of items on an instrument. Specifically, Cortina (1993) reported that Cronbach's alpha was underestimated on instruments with less than 19 items, as "the relationship between number of items and alpha is curvilinear (Komorita & Graham, 1965) and begins to level off before the number of items reaches 19" (p. 101). Since the index of psychosocial

functioning consisted of two items, its internal consistency reliability was likely underestimated. For instance, if an instrument with 10 items had an average inter-item correlation similar to that observed between the two items on the index of psychosocial functioning (i.e., $r = 0.33$), that instrument's internal consistency reliability would be above 0.80 (Schmitt, 1996).

Third, despite the existence of single criterions for internal consistency reliability (e.g., 0.80, Burlingame, Lambert, Reisinger, Neff, & Mosier, 1995; 0.70, Schmitt, 1996), it may be the case that a single criterion is inadequate when determining a measure's internal consistency reliability. In fact, an alpha coefficient as low as 0.50 can be acceptable when consideration is given to the degree of variation in the data set, number of items on the instrument, and "other desirable qualities, such as meaningful content coverage of some domain and reasonable unidimensionality" (p. 352; Schmitt, 1996; Cortina, 1993). Therefore, instead of using a single criterion, the acceptability of a measure's internal consistency reliability might best be determined on a case by case basis and interpreted in the context of other psychometric properties. After all, while Cronbach's alpha does say something about an instrument's unidimensionality in measuring a single construct, it says nothing about what that construct might be. An evaluation of some form of construct validity is needed to determine what an instrument is reliably measuring.

Convergent Validity of the Index of Psychosocial Functioning

The index of psychosocial functioning showed convergent validity by significantly correlating with other commonly used measures of psychosocial functioning

(e.g., DYS, GAF, and Q-LES-Q). These correlations of convergent validity were lowest during the month of diagnostic evaluations (*r* ranged from -0.34 to -0.63) and highest during the last month of A-CT (*r* ranged from -0.54 to -0.82). The noted increase in magnitude of correlation coefficients was likely due to the increase in variability in the dataset during the acute-phase. For instance, while all patients at the beginning of the acute-phase met criteria for MDD, all patients at the end of the acute-phase did not, with some showing few residual depressive symptoms, some showing no depressive symptoms, and some still meeting criteria for MDD. Of note, correlations between the index of psychosocial functioning and DYS, GAF, and Q-LES-Q were moderate to large in size, suggesting there was significant agreement but not enough to indicate measures were redundant. Also, Weissman et al (2001) and Vittengl et al. (2008) corroborated these results, showing the RIFT and SAS-SR, instruments that together form the index of psychosocial functioning, demonstrated convergent validity with the DYS, GAF, SASS, SF-36 Mental Component, and Inventory of Interpersonal Problems (Horowitz, Rosenberg, Baer, Ureno, & Villasenor, 1988).

Discriminant Validity of the Indices of Psychosocial Functioning and Depressive Symptom Severity

The indices of psychosocial functioning (comprised of RIFT and SAS-SR total scores) and depressive symptom severity (comprised of BDI, HRSD-17, and IDS-SR total scores) did not demonstrate discriminant validity in relation to each other. Specifically, correlations between indices ranged from 0.61 to 0.77 during A-CT. Also, scores on the index of psychosocial functioning highly correlated with total scores from

the BDI, HRSD-17, and IDS-SR, with *r* ranging from 0.59 to 0.77 during A-CT. All correlations of discriminant validity were large in size (Cohen, 1988), suggesting there was significant overlap in constructs.

This finding raised two important issues. First, the lack of discriminant validity between the indices in this study mirrored previous research showing the constructs of psychosocial functioning and depressive symptom severity were highly correlated (e.g., Judd, Akiskal, et al., 2000; Lepine et al., 1997; Mulder et al., 2003; Pedersen et al, 2002; Revicki et al., 1992). As such, the lack of discriminant validity between indices in this study may reflect an overarching limitation of existing instrumentation in the field more so than a limitation in measurement specific to this study (see p. 218 for a discussion of how the overlap between constructs represents a study limitation).

As a second and related issue, it could be that some of the inter-correlation between indices in this study is due to shared item content. For instance, Ro, Clark, Vittengl, and Jarrett (2007) suggested that items on the SAS-SR, a measure included on the index of psychosocial functioning, were confounded by their use of terms expressing negative affect when querying domains of psychosocial functioning (e.g., "Have you been *ashamed* [italics added] of how you do your work in the past two weeks?"). As a result, individual's self-reported levels of psychosocial functioning on the SAS-SR could also inadvertently reflect levels of their depressive symptomatology (e.g., depressed mood, excessive guilt or worthlessness, etc.).

Index Sensitivity to Pre- to Post-A-CT Changes

The indices of psychosocial functioning and depressive symptom severity showed they were sensitive enough to detect pre- to post-A-CT changes, as evidenced by their large respective standardized mean gain effect sizes of 1.54 and 2.37. These effect sizes paralleled those reported in RCTs of A-CT investigating changes in psychosocial functioning (e.g., Dunn et al., 2007) and depressive symptom severity (e.g., Bagby et al., 2008; Dimidjian et al., 2006; Elkin et al., 1989; Evans 1992; Hollon et al., 2005; Shea 1992; Thompson et al., 1987). Consequently, it could be inferred that the indices in this study's primary aim adequately reflected changes occurring in psychosocial functioning and depressive symptom severity during A-CT.

Summary of Psychometric Quality of Indices

Taken as a whole, results from this preliminary aim supported the psychometric quality of the index of psychosocial functioning and replicated past findings with the index of depressive symptom severity. Specifically, concerns over the internal consistency reliability of the index of psychosocial functioning were alleviated by: a) a review of factors that influenced this psychometric property (e.g., number of items) and b) the moderate to high convergent validity of the index with other commonly used measures of psychosocial functioning. Also, the indices' low discriminant validity did not warrant their removal from the primary aim, since this finding replicated previous research in illustrating a limitation of the field and limited measurement options available at the time this study was designed. After all, the primary aim in this study was designed to account for this reported limitation in the field, by covarying out the shared variability

between constructs at treatment baseline and the beginning, middle, and end of A-CT. Lastly, the indices of psychosocial functioning and depressive symptom severity were found to be sensitive to pre- to post-A-CT changes in their respective constructs. As such, these analyses substantiated the use of the indices of psychosocial functioning and depressive symptom severity in the primary aim.

Early Responders Report Significantly less Impairment in Psychosocial Functioning and Depressive Symptom Severity than Late Responders during Acute-Phase Cognitive Therapy

This study was the first to show early responders (i.e., ≥ 40% reduction in HRSD-17 scores by session eight of A-CT compared to treatment baseline) had less impairment in psychosocial functioning than late responders (i.e., < 40% reduction in HRSD-17 scores by session eight of A-CT compared to treatment baseline) at treatment baseline and the beginning, middle, and end of A-CT. The fact that this study provided late responders with more time in treatment (e.g., four additional sessions) to improve their response further highlighted this difference. Of note, research investigating pharmacological (Aberg-Wistedt, Agren, Ekselius, Bengtsson, & Akerblad, 2000; Koran et al., 1995; Mulder, Joyce, Frampton, Luty, & Sullivan, 2006; Papakostas, Petersen, Denninger, et al., 2004; Pollock, Perel, Kupfer, Bowler, & Miewald, 1993; Trivedi, Morris, Grannemann, & Mahadi, 2005) and other psychosocial treatments (Arnow et al., 2007) for depression substantiated this finding, continually showing those who responded early to treatment had favorable acute-phase outcomes compared to those who did not.

Past research has not yet determined if early responders have better acute-phase outcomes than late responders because they: a) start off with less impairment; b) are more reactive to treatment; c) or both. In this study, early responders started A-CT with less impairment in psychosocial functioning and depressive symptom severity than late responders, and they continued with less impairment throughout treatment. Therefore, results from this study would suggest that early responders had better outcomes to A-CT than late responders because they began treatment healthier.

Significant Reductions in Psychosocial Functioning and Depressive Symptom Severity during Acute-Phase Cognitive Therapy

This study showed depressive symptom severity and psychosocial functioning significantly improved during A-CT. Reductions in both constructs were most likely attributable to time spent in A-CT, rather than possible confounds or regression to the mean. This inference was based on three findings. First, potential confounding variables (i.e., "early vs. late response to A-CT") accounted for ≤ 2% of the variability in change on the indices of psychosocial functioning and depressive symptom severity during A-CT. Second, "time in A-CT" (i.e., indices measured at treatment baseline and the beginning, middle, and end of A-CT) accounted for 46% and 22% of the variability in improvements on the indices of psychosocial functioning and depressive symptom severity, respectively. Thirdly, effect sizes representing pre- to post-A-CT changes on the indices of psychosocial functioning and depressive symptom severity and response rates replicated those reported in RCTs investigating the impact of A-CT on these

constructs (e.g., Dunn et al., 2007; Elkin et al., 1989; Hollon et al., 2005; Thompson et al., 1987). Based on these points, it was inferred that the reductions observed in this study in psychosocial functioning and depressive symptom severity were comparable to exposure to A-CT and not likely due to regression to the mean, or some other confounding variable.

Also of interest, this study provided information regarding the impact of additional, non-protocol treatments on psychosocial functioning and depressive symptom severity in patients receiving A-CT for MDD. Results showed patients most often were non-compliant with their agreement to avoid non-protocol mood altering treatment when they suffered insomnia, as the most frequent non-protocol treatment was sleep-aids. While there was no statistically significant impact on either construct, a trend was noted such that patients who received additional, non-protocol treatment reported more depressive symptom severity (M = 30.25, SD = 16.95) than those who did not (M = 21.52, SD = 13.52) at the post-A-CT blind evaluation (F [3, 719] = 2.19, p = 0.09). This finding may suggest that patients who experienced less than desirable outcomes during A-CT were more likely to seek additional treatment outside study protocol than those who achieved desired outcomes. Whether or not this was the case, however, studies should track and report non-protocol treatment obtained during clinical trials for depression to better understand their cause and impact on protocol treatment outcomes.

Changes in Psychosocial Functioning during A-CT in Relation to Changes in Depressive Symptom Severity

By controlling for shared sources of error variance across repeated measurements and covariation between constructs, the structural equation model in this study sought to isolate variance unique to indices of psychosocial functioning and depressive symptom severity. When this variance was isolated and baseline levels of psychosocial functioning were controlled, the model showed psychosocial functioning at the first month of A-CT partially mediated changes in depressive symptom severity from baseline to week seven of A-CT. Specifically, a one point increase on the baseline index of depressive symptom severity led to a 0.09 point increase in the index of depressive symptom severity at week seven of A-CT, via its prior effect on psychosocial functioning during the first four weeks of A-CT. This pattern of results suggested that improvements in psychosocial functioning at the beginning of A-CT helped explain reductions in depressive symptom severity during the first seven weeks of A-CT.

The above mediating impact of psychosocial functioning was partial because significant changes occurred in depressive symptom severity during A-CT independent of changes in psychosocial functioning. These independent changes were reflected in the mediating impact of depressive symptom severity at week one of A-CT on changes in the same construct from baseline to week seven of A-CT. In terms of the index of depressive symptom severity, a one point increase at baseline led to a 0.24 point increase at week seven of A-CT, via prior effects on the index at week one of A-CT. Together, the indices of depressive symptom severity and psychosocial functioning during the first four weeks

of A-CT completely mediated changes in depressive symptom severity from treatment baseline to week seven of A-CT.

After week seven of A-CT, significant, independent changes occurred in each construct, and psychosocial functioning significantly predicted subsequent depressive symptom severity. Specifically, a one point increase on the index of psychosocial functioning at the second month of A-CT led to a 0.40 point increase in post-A-CT scores on the index of depressive symptom severity. This meant that greater impairment in psychosocial functioning at the middle of the acute-phase led to worse depressive symptom severity immediately after A-CT. On the other hand, depressive symptom severity at week seven of A-CT did not significantly predict subsequent psychosocial functioning. In fact, only one path showed that depressive symptom severity predicted subsequent psychosocial functioning during A-CT (i.e., index of depressive symptom severity at diagnostic evaluation → index of psychosocial functioning at month one of A-CT; see Figure 13).

These findings were similar to those reported by Vittengl et al. (2008) who studied the extent to which psychosocial functioning (measured by the RIFT) predicted depressive symptom severity (measured by the LIFE-PSR) after A-CT in responders. For their study, responders to A-CT for MDD were randomized to either eight months of continuation Cognitive Therapy or assessment-only control and followed for an additional 16 months. After controlling for past psychosocial functioning, they found that depressive symptom severity did not predict future psychosocial functioning. However, they reported that psychosocial functioning predicted future depressive symptom severity, when past depressive symptoms were controlled. Taken together, like

the current study, Vittengl et al. (2008) showed psychosocial functioning influenced later depressive symptom severity and not vice versa.

However, four investigations of acute-phase treatment disagreed with this study's results. For example, the same research group found that changes in depressive symptom severity completely accounted for changes in psychosocial functioning in a sample diagnosed with MDD receiving A-CT (Vittengl et al., 2004), while changes in psychosocial functioning partially accounted for changes in depressive symptom severity. In their analyses, Vittengl et al. operationalized psychosocial functioning, using scores from the SAS-SR and DYS, and depressive symptoms, using an index that combined total scores from the BDI, HRSD-17, and both versions of the IDS (i.e., self-report and clinician-rated). When measuring psychosocial functioning with the General Life Functioning Scale, Lenderking et al. (1999) also found changes in depressive symptom severity, measured by the MADRS, completely accounted for changes in psychosocial functioning during a seven week trial of venlafaxine. Furthermore, Hirschfeld et al. (2002) and Finkelstein et al. (1996) found changes in depressive symptom severity, measured by the HRSD-17, partially accounted for changes psychosocial functioning, measured by the SAS-SR and a homemade measure of work functioning, respectively. While Hirschfeld et al. investigated a 12 week trial of nefazodone and Cognitive Behavioral Analysis System of Psychotherapy, Finkelstein et al. examined the effects of a 12-week trial of sertraline and imipramine. In sum, this body of research suggested that changes in depressive symptom severity partially to completely accounted for changes in psychosocial functioning, while changes in psychosocial functioning only partially accounted for changes in depressive symptom severity.

Five reasons may account for this study's divergence from the four studies just reviewed. First, while this study used different raters (e.g., self-report and clinician) and types of measurement (e.g., cross-sectional and longitudinal), the four studies just reviewed only used self-report, cross-sectional measures to assess psychosocial functioning. It may be that clinician-rated, longitudinal measures, like the RIFT, tap into unique aspects of psychosocial functioning that alter the mediating relationship between psychosocial functioning and depressive symptom severity during A-CT (see secondary analysis on p. 174).

Second, when investigating the extent to which change in psychosocial functioning accounted for change in depressive symptom severity, or vice versa, the four studies just reviewed did not establish precedence of change (Finkelstein et al., 1996; Hirschfeld et al., 2002; Lenderking et al., 1999; Vittengl et al., 2004). Instead, these studies sought to account for pre- to post-treatment change in one construct with concurrent change in the other. As a result, even if changes in depressive symptom severity accounted for changes in psychosocial functioning, it could not be ruled out that the mediational relationship also existed when constructs were reversed (Kraemer et al., 2001). On the other hand, this study established temporal precedence by only examining mediators that occurred after treatment began and before the dependent variable was measured.

Third, as just mentioned, previous studies of the relationship between changes in psychosocial functioning and depressive symptom severity only examined acute-phase changes from pre- to post-treatment. The acute-phase in these studies ranged from seven (Finkelstein et al., 1996) to 12 weeks (Hirschfeld et al., 2002; Lenderking et al., 1999;

Vittengl et al., 2004) in length. Comparatively, this study evaluated change in psychosocial functioning and depressive symptom severity that occurred from treatment baseline to the beginning, middle, and end of A-CT. Specifically, psychosocial functioning was measured every month of the acute-phase, and depressive symptom severity was measured at week one and seven of A-CT, as well as at the post-A-CT blind evaluation. By splitting up the acute-phase into shorter time intervals, this study may have been more sensitive to changes in the reciprocal relationship between psychosocial functioning and depressive symptom severity (see secondary analysis on p. 191). For instance, this study found that change in psychosocial functioning only mediated change in depressive symptom severity between treatment baseline and week seven of A-CT.

Fourth, this study's results may differ from previous research due to the use of structural equation modeling. Instead of structural equation modeling, the four studies reviewed above used linear regression to investigate change in psychosocial functioning in relation to change in depressive symptom severity (Finkelstein et al., 1996; Hirschfeld et al., 2002; Lenderking et al., 1999; Vittengl et al., 2004). According to Kline (2005), mediational analyses based on linear regression were limited by an inability to enter variables as both predictor and criterion in the same analysis or control for covariation between variables across data sets. As a result, this study may have detected a different mediational relationship between psychosocial functioning and depressive symptom severity because structural equation modeling did not have the above limitations (see secondary analysis on p. 196).

Finally, it may be that a yet to be identified mechanism was at work where early improvements in psychosocial functioning influenced depressive symptom severity at the

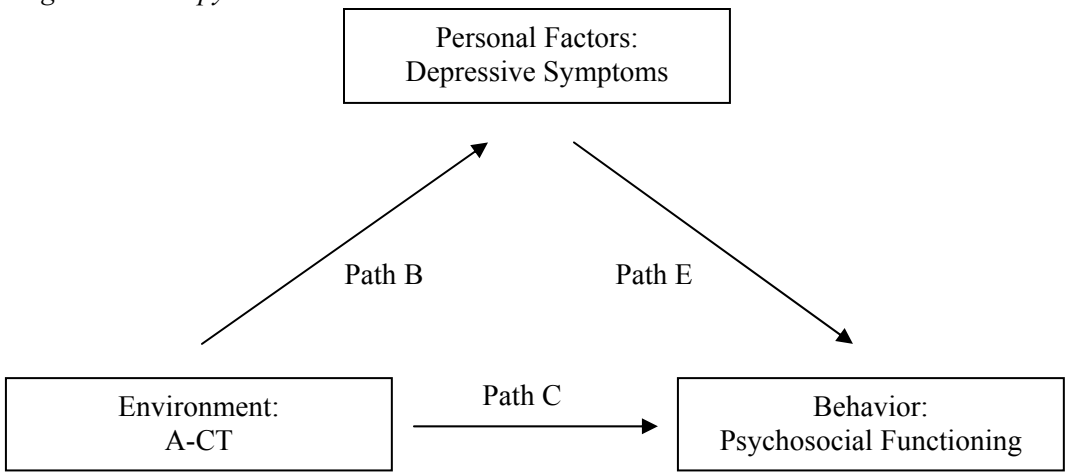
beginning of the acute-phase. When comparing the change slopes of the indices of psychosocial functioning and depressive symptom severity (see Figure 6 on p. 137), it appeared that psychosocial functioning improved less and more gradually during A-CT than depressive symptom severity. Even though psychosocial functioning improved more slowly than depressive symptom severity, these early improvements in psychosocial functioning, or maybe just the initial mobilization of resources to improve psychosocial functioning, may have served an important function in the alleviation of depressive symptom severity.

Study Implications

Results from this study may have implications on how the field understands and measures change in psychosocial functioning during A-CT for MDD. This study primarily tested a theoretical framework based on social cognitive theory that conceptualized change in psychosocial functioning during A-CT. In terms of this theoretical framework, it was originally hypothesized that changes in personal factors (i.e., depressive symptom severity) would partially mediate changes in behavior (i.e., psychosocial functioning) when people were exposed to A-CT (i.e., environmental stimuli) for MDD (see Path E in Figure 20). It was also predicted that significant, independent changes would occur in each construct during A-CT (see Paths B and C in Figure 20).

Figure 20

Hypothesized Mediating Relationship between Psychosocial Functioning and Depressive Symptom Severity in Terms of Triadic Reciprocal Causation during Acute-Phase Cognitive Therapy

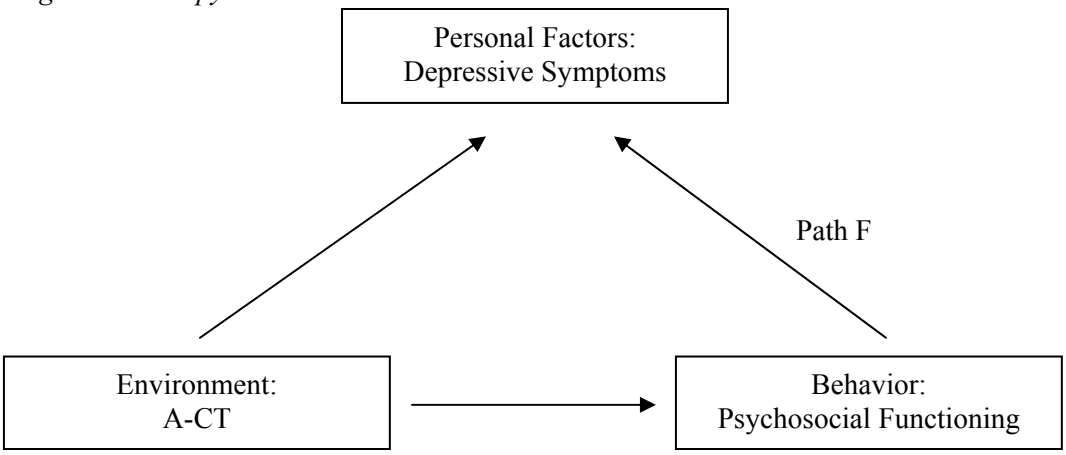


Note. Adapted from Bandura (1986); A-CT = Acute-Phase Cognitive Therapy.

While each construct did significantly change independently of each other during A-CT (see Path B and C in Figure 20), study results suggested that when people with MDD were exposed to A-CT, change in their behavior (i.e., psychosocial functioning) partially mediated change in personal factors (i.e., depressive symptom severity) (see Path F in Figure 21).

Figure 21

Observed Mediating Relationship between Psychosocial Functioning and Depressive Symptom Severity in Terms of Triadic Reciprocal Causation during Acute-Phase Cognitive Therapy



Note. Adapted from Bandura (1986); A-CT = Acute-Phase Cognitive Therapy.

Consequently, even though psychosocial functioning improved more slowly during the acute-phase than depressive symptom severity (e.g., Bothwell et al., 1977; Mintz et al., 1992), early changes in functioning appeared to have some benefit on subsequent reductions in depressive symptom severity.

This finding would be consistent with research showing that the specific behavioral components of A-CT effectively reduced depression symptom severity (Follette & Greenberg, 2005). Regarding this issue, Beck et al. (1979) wrote that:

> In the early stages of cognitive therapy and particularly with the more severely depressed patients, it is often necessary for the therapist to concentrate on restoring the patient's functioning to the premorbid level… The rationale for this

> approach is based on the clinical observation that the severely depressed patient, and often the important people in his life, believe that he is no longer capable of carrying out the typical functions expected in his role as a student, wage earner, homemaker, spouse, parent, etc. Furthermore, the patients can see no hope of gaining satisfaction from those activities… (p. 117).

Clark et al. (1999) went on to say that "a shift from a negative to euthymic mood state may not be possible" (p. 91) until schemas were activated that involved a patient's ability to engage in, and benefit from, interpersonal and occupational roles. Therefore, by focusing on behavior change early in the acute-phase, cognitive behavioral therapists expected to improve depressive symptom severity by identifying and challenging underlying maladaptive patterns of cognition, increasing access to reinforcement, and reducing exposure to punishment (Beck et al., 1979; Lewinsohn et al., 1985).

Theoretically, the above finding (see Figure 15) would also be consistent with interventions that focused on improving patients' interpersonal functioning, or social milieu, in order to reduce depressive symptomatology (e.g., Interpersonal Psychotherapy, Interpersonal and Social Rhythm Therapy, Family-Focused Therapy, etc.; Miklowitz et al., 2007; Weissman, Markowitz, & Klerman, 2000). For example, Interpersonal Psychotherapists claim that "change and improvement in depressive symptoms occur through working on mastery and competence in the social sphere" (p. 44; Crowe & Luty, 2005). Therefore, it could be that this study's conceptualization of change in psychosocial functioning in relation to depressive symptom severity (see Figure 21) is robust and applicable to other psychosocial interventions. Yet, more research is needed before this claim can be validated.

In addition to furthering the field's conceptual understanding of how A-CT treats MDD, this study may also have practical implications for the measurement of psychosocial functioning during depression-specific, acute-phase treatment. For example, results suggested that measures of psychosocial functioning (especially the SAS-SR as seen in Appendix G) correlated highly with measures of depressive symptom severity. In fact, it may be difficult to tease apart variance unique to each construct due to the overlap between measures in content and shared method variance. This study's results therefore advocated the revision of existing instruments, or development of new instruments, to more accurately assess content specific to psychosocial functioning.

Study Limitations

This study was limited by the lack of discriminant validity between constructs, assessment strategy, study design, restrictions in generalizability, and use of correlational data. First, as just mentioned, the indices of psychosocial functioning and depressive symptom severity highly correlated, with r ranging from 0.61 to 0.77. This lack of discriminant validity likely reflected imperfections in technology related to overlap in item content (Ro et al., 2007). What is still unclear, however, is the degree to which psychosocial functioning and depressive symptom severity actually overlap. For example, some concepts may legitimately pertain to both constructs (e.g., satisfaction with social role performance may depend on both mood and objective functioning), while others may not (e.g., suicidal ideation). As a result, it is also unclear how much covariation between psychosocial functioning and depressive symptom severity should

be controlled. If, for instance, a structural equation model controlled for covariation between constructs that actually existed, the external validity of the mediating relationship detected by the model would be limited. Consequently, until the constructs in this study are better understood and technology improved to increase discriminant validity, the mediating relationship in this study should be viewed as exploratory.

Second, issues may arise regarding this study's assessment strategy, particularly its use of indices, shared method variance, use of the RIFT in a retrospective manner, and lack of a longitudinal measure of depressive symptom severity. While indices were created to reduce spurious results due to differences between instruments, the use of indices may also have limited the degree to which clinical settings can replicate this study, as resources, providers, and staff are typically overburdened in these settings. Furthermore, since the same instruments were included in the indices of depressive symptom severity and psychosocial functioning across all time points, some "shared method variance may be inextricably entwined with the constructs" (p. 569; Cole et al., 2003). Consequently, despite efforts in the primary aim to account for shared method variance (i.e., use of indices that combine data from multiple informants, addition of paths in the structural equation model that control for shared error variance across repeated measurements), this study may be more susceptible to Type I errors than if steps had been taken to measure each construct differently at each time point.

In addition, because the RIFT was used in a retrospective manner, clinicians relied heavily on patients' ability to recall events from their past when completing it, thereby potentially introducing mood-congruent memory biases in the index of psychosocial functioning (Barry et al., 2004). Lastly, results from the primary aim might

have been different if the index of depressive symptom severity included a longitudinal measure. As it was, only the index of psychosocial functioning included a measure that accounted for changes during the entire acute-phase. Therefore, this study's findings may be attributable to differences in cross-sectional vs. longitudinal measurements rather than true treatment effects across constructs. Future research might improve on this study's assessment strategy, by adapting it to clinical settings (e.g., remove one self-report measure of depressive symptom severity, have staff administer clinician-rated instruments instead of providers, etc.), using different instruments to measure constructs at different time points, using the RIFT prospectively, and including a longitudinal measure of depressive symptom severity.

Third, this study's design was limited due to its lack of random assignment and a control group. Ideally, a randomized controlled trial would be implemented to control for the impact of extraneous factors (e.g., patient maturation, environmental stimuli, etc.) on changes in psychosocial functioning and depressive symptom severity during A-CT. With such a study design, statistically significant differences between treatment groups could be attributed to A-CT.

Instead, this study adopted the pre- to post-treatment design used during the acute-phase in the original NIMH funded studies (i.e., "Prophylactic Cognitive Therapy of Depression" [R01 MH 58397 and 58396] and "Are Cognitive Therapy's Antidepressant Effects Durable?" [R01 MH 69619 and 69618]). To make up for the lack of random assignment and a control group, the current study sought to attribute changes in psychosocial functioning and depressive symptom severity to exposure to A-CT by: a) statistically controlling for unexplained sources of variation and b) comparing the

magnitude of changes and response rates to those from RCTs. Despite these efforts and past RCTs showing A-CT for MDD effectively reduced depressive symptom severity and impairment in psychosocial functioning (e.g., Gloaguen et al., 1998; Imber et al., 1990; Vittengl et al., 2007), this study's statistical analyses inherently left open the question, to some degree, "What caused the changes that occurred during A-CT?"

Fourth, study conclusions may have limited generalizability due to sample demographics and treatment specificity. Despite targeting ethnic/racial minorities with specific recruitment strategies, this study slightly over- and undersampled Whites and ethnic minorities, respectively, compared to current racial trends in the U.S. (U.S. Census Bureau, 2002). This was largely due to the low number of ethnic minority patients at the WPIC site. Also, since all patients in this study met criteria for recurrent MDD, results may not generalize to populations with other diagnoses.

Furthermore, this study's conclusions may only generalize to patients undergoing A-CT, as the change observed in this study in psychosocial functioning and depressive symptom severity may represent treatment effects unique to A-CT. Researchers are currently debating this issue. On one hand, researchers claim that despite the existence of hundreds of therapeutic techniques and potential mechanisms of change (Froyd, Lambert, & Froyd, 1996; Hill, Nutt, & Jackson, 1994; Kazdin, 1986), little evidence exists to show one is more effective than the others in terms of treatment outcome (e.g., Lambert et al., 2004; Luyten et al., 2006; Wampold et al., 1997). Even for empirically supported treatments (ESTs) of depression, "there is relatively little research evidence indicating that the observed response to psychosocial ESTs for depression is primarily due to particular theoretical factors or specific therapeutic techniques" (p. 989; Luyten et al.). In

fact, researchers have found that factors common to all psychosocial treatments (e.g., therapeutic alliance) have twice the impact on outcomes (e.g., psychosocial functioning, symptom severity, etc.) than factors specific to treatment strategies (Blatt & Zuroff, 2005; Lambert & Barley, 2002). According to this research, the patterns of change observed in this study may generalize to other psychosocial treatments besides A-CT.

On the other hand, researchers have also argued that psychosocial treatments differ in their ability to treat specific disorders. For instance, Gloaguen et al. (1998) claimed cognitive therapy was "superior" to therapies that did not employ cognitive or behavioral components in the treatment of depression (p. 69). Leichsenring and Rabung (2008) reported that long-term psychodynamic psychotherapy was superior to other psychotherapies (e.g., CT, dialectical-behavioral therapy, family therapy, supportive therapy, short-term psychodynamic therapy) when treating personality disorders and chronic mental illness. Also, Roth and Fonagy (2004) pointed out that research supported behavioral therapies as the most effective psychosocial treatment for specific phobias and obsessive-compulsive disorder. According to this sampling of studies, treatment outcomes appear to differ across mental illnesses and psychosocial interventions. As a result, it would be best for future researchers to clarify the extent to which this study's conclusions and conceptualization of psychosocial functioning generalize to other modalities of treatment.

Finally, structural equation modeling is often used to estimate presumed causal relationships. However, since the model in this study analyzed covariations in the data set, results pertaining to the meditating relationship of psychosocial functioning on depressive symptom severity could not be interpreted as causal. Also, a mediating

relationship can only imply causation if: a) variables show temporal precedence, b) causal pathways between variables are correctly specified, and c) all potential causal variables are accounted for (Kline, 2005). While this study established temporal precedence of change, only replication of study results could determine the extent to which variables were correctly specified in the structural equation model. Also, it is likely that all variables influencing change in psychosocial functioning and depressive symptom severity during A-CT were not included in the structural equation model. These points add further caution to claims of causation based on this study's results.

Future Research

Future research should first focus on clarifying the extent to which study conclusions generalize to more ethnically diverse patient populations and treatment modalities. For instance, it may be that this study's conceptualization of psychosocial functioning does not apply to pharmacological treatments or psychosocial interventions without a behavioral component. In these treatments, it may be that changes in depressive symptom severity partially mediate changes in psychosocial functioning, as was originally hypothesized.

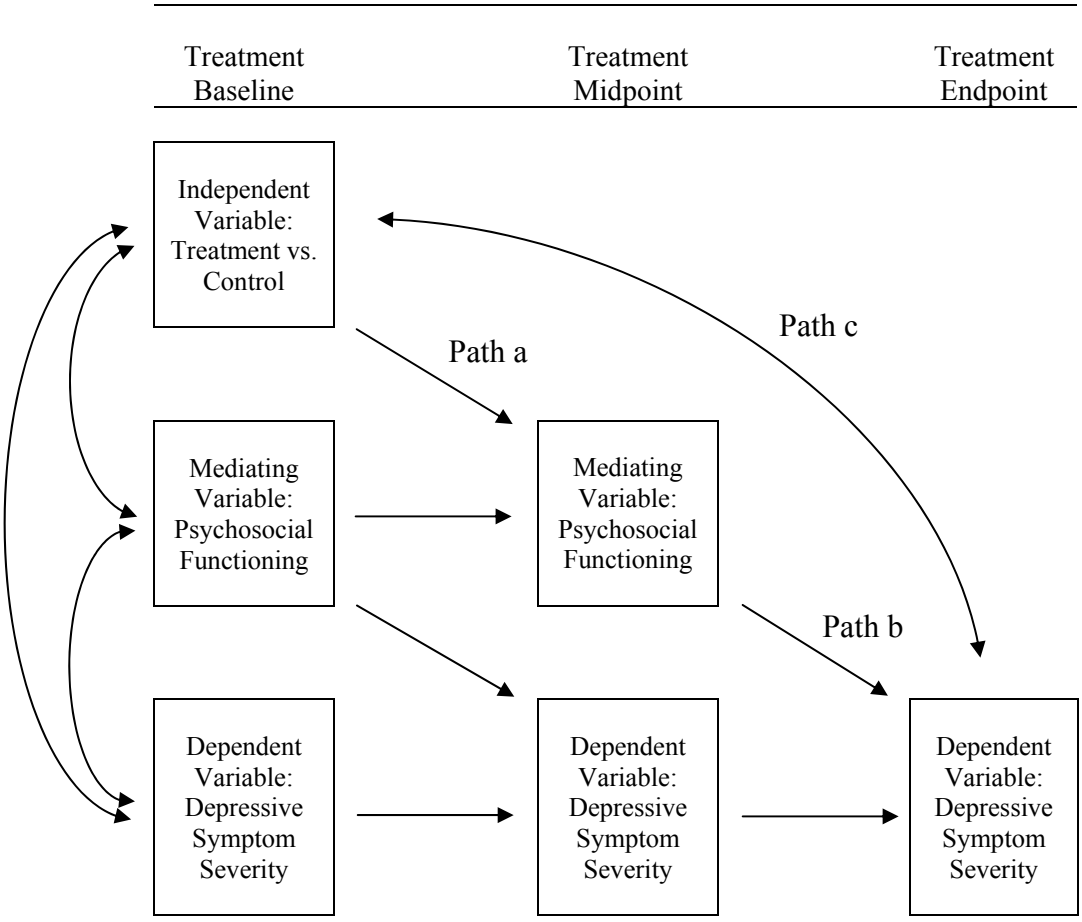
Future research is also needed to clarify the extent to which this previously unreported pattern of change (i.e., change in psychosocial functioning partially mediated change in depressive symptom severity during A-CT) is attributable to the use of clinician-rated and/or longitudinal measures of psychosocial functioning. It may be that clinician-rated instruments tap into a unique area of psychosocial functioning that self-

report instruments do not address. For instance, self-reported changes in psychosocial functioning "may represent subjective experiences rather than independently observable behavioral change", which is captured by clinician-rated measures of the construct (p. 655; Vittengl et al., 2004). It may also be the case that longitudinal measures are more sensitive to gradual changes that occur in psychosocial functioning during acute-phase treatment than cross-sectional measures. To investigate this issue, future researchers could replicated the primary aim, using different types of instruments (i.e., cross-sectional vs. longitudinal) with different raters (i.e., self-report vs. clinician-rated) to assess psychosocial functioning and depressive symptom severity.

As mentioned before, replication of this study's hypotheses with a study design that randomly assigns patients to treatment and control groups is also warranted. Within such a study design, researchers could fit the structural equation model with longitudinal data that accounts for: a) levels of the independent variable measured at time one (i.e., treatment vs. control group), b) the mediator measured at time one and two, and c) the dependent variable measured at time one, two, and three (Cole et al., 2003). If researchers then want to test the mediating impact of psychosocial functioning on depressive symptom severity, their structural equation model would look like the model in Figure 22. The model in Figure 22 would replicate study findings if the indirect effect of treatment on depressive symptom severity through the mediator psychosocial functioning (represented by paths a and b) is significant, and the direct effect of treatment on depressive symptom severity is not (represented by path c).

Figure 22

Structural Equation Model Fit with Longitudinal Data Showing Mediating Relationships between Psychosocial Functioning and Depressive Symptom Severity in a Randomized Controlled Trial



Note. Graph was adapted from Cole et al. (2003)

Since the measures of psychosocial functioning reviewed in this study (DYS, GAF, RIFT, SAS-SR, and Q-LES-Q) show generally acceptable internal consistency reliability, convergent and discriminant validity, and sensitivity to changes during A-CT

(see Appendix G for additional information), their continued use in clinical trials for depression-specific treatment is justified. That being said, however, ongoing refinements are warranted to improve their psychometric quality, with special emphasis on their discriminant validity with measures of depressive symptom severity. Toward this end, items can be developed that measure satisfaction and objective performance in occupational, interpersonal, and recreational roles, while avoiding references to depressive symptomatology (e.g., mood, energy level, etc.; Ro et al., 2007). Also, items could be added to instruments (e.g., items addressing satisfaction with role performance added to SAS-SR) so that each domain of psychosocial functioning is adequately covered. Such improvements in technology may reduce inter-correlation between measures of psychosocial functioning and depressive symptom severity, thereby improving researchers' ability to investigate each construct during diagnostic and treatment outcome evaluations.

Finally, it appears that psychosocial functioning can be understood in terms of distinct domains (e.g., occupational, interpersonal, and recreational functioning). Moreover, this and past studies suggest that the specific domains of psychosocial functioning respond differently to A-CT (Vittengl et al., 2004). For instance, the DYS, which measures dyadic functioning, changes less during A-CT than measures of overall psychosocial functioning (e.g., SAS-SR; see Chapter 3 or Appendix G for more information). It may be that only those domains of psychosocial functioning that change fastest influence change in depressive symptom severity during A-CT. Future research could investigate potential differential treatment effects between domains of psychosocial functioning, by examining the relationship between changes in each domain and changes

in depressive symptom severity. Specifically, researchers could seek to replicate the primary aim with indices that measure a single domain of psychosocial functioning (e.g., work functioning), rather than overall psychosocial functioning.

Summary

Acute-phase CT, like other psychotherapeutic interventions, is a complex intervention with multiple potential mechanisms of change (e.g., environmental, biological, cognitive, etc.; Garratt et al., 2007; Whisman, 1993). While the field has developed some understanding of how A-CT reduces impairments in psychosocial functioning in relation to depressive symptom severity (e.g., Hirschfeld et al., 2002; Vittengl et al., 2004), this area of research is still in its infancy. By disentangling the sequence of change in psychosocial functioning and depressive symptom severity, this study helps researchers and clinicians come one step closer to understanding how A-CT treats the impairment in psychosocial functioning associated with MDD. Results suggest that: a) changes in psychosocial functioning and depressive symptom severity occur independently of each other throughout A-CT, b) changes in psychosocial functioning during the first month of A-CT partially mediate changes in depressive symptom severity from treatment baseline to week seven of A-CT, and c) psychosocial functioning at week seven of A-CT significantly predicts subsequent depressive symptom severity. Study results are limited by the lack of random assignment and a control group, as well as by limitations in existing instruments when measuring the construct of psychosocial functioning. Future research should seek to replicate these results in more diverse patient

populations and treatment modalities. Also, short-term efforts on instrument revision and/or development may improve the degree to which researchers can isolate and evaluate the role of psychosocial functioning in diagnostic and treatment outcome evaluations.

APPENDIX A

Dunn, T. W., & Jarrett, R. B. (in press). Psychosocial Functioning in Depression. In R. Ingram (Ed.), ***The International Encyclopedia of Depression*****.**

Running Head: PSYCHOSOCIAL FUNCTIONING

Psychosocial Functioning in Depression

Todd W. Dunn and Robin B. Jarrett

The University of Texas Southwestern Medical Center at Dallas

Department of Psychiatry

The International Encyclopedia of Depression

Submitted: January 31, 2008

Address correspondence to: Todd W. Dunn or Robin B. Jarrett, The University of Texas Southwestern Medical Center at Dallas, Department of Psychiatry, 5323 Harry Hines Blvd., Dallas TX 75390-9149 (phone: 214-648-5345, e-mail: Todd.Dunn@UTSouthwestern.edu or Robin.Jarrett@UTSouthwestern.edu)

Psychosocial Functioning in Depression

Recent estimates suggest 1 out of 6 Americans will suffer from depression during their lifetime. Since the 1960s, researchers have consistently shown that individuals diagnosed with depression have greater impairment in all areas of psychosocial functioning than people who are not depressed. In fact, to be diagnosed with Major Depressive Disorder (MDD) or Dysthymia, the persistent, associated depressive symptoms listed in the criteria must cause "significant distress or impairment in social, occupational, or other important areas of functioning"(American Psychiatric Association, 2000). In other words, impairment in psychosocial functioning is "built into" the diagnosis of depression. The severity of this psychosocial impairment rivals that found in such chronic diseases as cardiovascular disease, cancer, and arthritis (Hays, Wells, Sherbourne, Rogers, & Spritzer, 1995) and represents the "leading cause of disease-related disability among women in the world today" (p. 5; Kessler, 2003).

As suggested above, impairment in psychosocial functioning is linked to changes in the course of illness, such as onset and persistence of depressive symptoms, poor treatment response, and higher rates of relapse and recurrence. In addition, researchers have shown that the impairment in psychosocial functioning associated with depression can be stable over time, especially in patients with depression who suffer treatment-resistance or psychosis. Estimates of this impairment increase further in patients with depression that also have non-psychiatric medical disorders.

Definition of Psychosocial Functioning

Given the prominence of psychosocial functioning in the experience and diagnosis of depression, it is surprising that the role and measurement of psychosocial

functioning is not more frequently studied. At present, psychosocial functioning can be thought of as person's performance in and satisfaction with his or her occupational, social-interpersonal, and recreational roles (Hirschfeld et al., 2000). In other words, psychosocial functioning is based on both objective and subjective assessments of role performance. Objective estimates of role performance, involves someone besides the person who is depressed describing or quantifying the behaviors which fulfill the depressed person's social obligations. Subjective estimates of satisfaction involve the depressed person quantifying the extent to which fulfilling these roles meets personally relevant standards, which include both satisfaction with the role itself and the person's performance within the role.

Table 1 provides a heuristic for conceptualizing psychosocial functioning. Social-interpersonal functioning can be viewed as fulfilling or performing within such social roles as a partner, lover, significant other, parent, family member, extended relative, friend, and/or neighbor. Within these roles, consideration is given to the frequency, quality, and content of relationships, including sexual functioning. With occupational functioning, a person rates his or her performance at work, school, and/or home in such terms as investment, workload, or productivity. Finally, to rate their recreational functioning, people report on the quantity and quality of "leisure" activities and hobbies (e.g., sports, gardening, reading, watching TV, going to movies, church, etc.).

Social roles and evaluation of functioning within each role can overlap and may conflict. For example, a father who plays with his children after work can be functioning

in both the recreational and social-interpersonal domains. Likewise, an employee who aggressively (but rudely) promotes her own agenda items in a business conference may perceive herself as functioning well within the occupational domain, while her colleagues perceive her as having poor social and interpersonal skills.

Table A1

Measuring Psychosocial Functioning using both Objective and Subjective Perspectives

Measurement of Psychosocial Functioning

Objective assessment of performance in each role	Interpersonal Functioning with:	Occupational Functioning at:	Recreational Functioning in:
	Spouse/Partner	Work	Hobbies
	Parents	School	Exercise
	Children	Home	Religious service
	Siblings		Community service
	Extended family		Other recreational activities
	Neighbors		
	Friends		

← Satisfaction with content and performance in each role →

Findings relevant to each domain of psychosocial functioning are abstracted below.

Social-Interpersonal Functioning

Depressed people have more difficulty in their relationships with spouses, co-workers, family, and friends than people who are not depressed. Bothwell and Weissman (1977) reported that this impairment in intimate relationships can persist for over four years, even after the complete remission of depressive symptoms. Interestingly, family

members of depressed patients also endorse greater impairment in psychosocial functioning than normative samples, suggesting that the family members might also be depressed or have suffered a loss in psychosocial functioning due to some yet to be identified mechanism.

Occupational Functioning

When considering occupational impairment, depression costs employers more than any other psychiatric disorder and accounts for more days missed at work than chronic medical conditions, such as diabetes, heart disease, hypertension, and lower back pain (Williams & Strasser, 1999). In fact, epidemiologists estimate 172 million days of work are lost each year in the U.S. due to depression (Dew, Bromet, Schulberg, Parkinson, & Curtis, 1991). Researchers also estimate that the U.S. loses $51.5 billion each year due to absenteeism and reduced productivity related to depression (Greenberg et al., 2003).

In addition, when depressed workers lose their employment, they have difficulty accessing health care services, as many only receive health care benefits through work. Of these individuals, only 33% go on to receive treatment, compared to 40% or 54% of individuals with depression who stay employed or are otherwise out of the work force, respectively (Greenberg et al., 2003). Even when treatment is obtained, however, improvements in occupational functioning take time and are "undone" with each recurrence of depressive symptoms (Mintz, Mintz, Arruda, & Hwang, 1992).

Recreational Functioning

Compared to social-interpersonal or occupational functioning, recreational functioning has received the least attention in the depression literature. Despite this,

researchers show individuals with depression report greater impairment in recreational functioning than those without depression. This impairment is most prevalent in individuals over the age of 60 (De Lisio et al., 1986). Like social-interpersonal and occupation functioning, impairment in recreational functioning is relatively stable and may persist even after individuals experience remission of depressive symptoms.

Treatment Effects on Psychosocial Functioning

Researchers have shown that both pharmacological and psychosocial interventions significantly reduce impairment in psychosocial functioning in depressed populations. The data suggest that exposure to pharmacotherapy, Interpersonal Psychotherapy, and Cognitive-Behavioral Therapy (CBT), is associated with improvements in social-interpersonal, occupational, and recreational functioning (e.g., Imber et al., 1990; Papakostas et al., 2004). These findings have been replicated with those receiving inpatient or outpatient treatment for acute or chronic depression, and with primary care, elderly, low income, and minority women populations. Of note, researchers suggest that psychosocial interventions involving both partners (e.g., Marital Behavioral Therapy) are most effective at reducing impairment in marital or sexual functioning (e.g., Beach & O'Leary, 1992).

Despite these encouraging results, however, the data are inconclusive regarding the extent to which depression-specific treatment restores or "creates" psychosocial functioning that falls within so-called "normal" ranges. In a 12-week trial of antidepressant medication, Miller et al. (1998) showed that chronically depressed outpatients had significantly more impairment in psychosocial functioning after treatment than normative samples. In contrast, Vittengl, Clark, & Jarrett (2004) reported that over

60% of outpatients with recurrent MDD returned to normative, "healthy" levels of psychosocial functioning after 12 to 14 weeks of CBT. Researchers have yet to determine the extent to which these disparate findings reflect the influence of different treatment modalities, measurement intervals, and/or effects or special characteristics of setting criteria for normality, distinct patient populations, or disease processes.

Researchers also remain uncertain about the timing and mechanisms of change in psychosocial functioning during treatment and across the course of illness. This may be due to the high correlation between psychosocial functioning and depressive symptom severity before, during, and after treatment. On one hand, improvements in psychosocial functioning may precede and facilitate reductions in depressive symptoms. If this hypothesis is substantiated, depression-specific interventions may more effectively reduce depressive symptoms by initially targeting and improving functional impairment. Alternatively, reductions in depressive symptoms may also account for improvements in psychosocial functioning, as shown by Vittengl et al. (2004). If research continues to support this second hypothesis, efforts to treat impairment in psychosocial functioning in individuals with MDD may first focus on reducing depressive symptomatology. With either pathway, a better understanding of the timing and process of change in psychosocial functioning during treatment will facilitate researchers' and clinicians' efforts to optimize targeted interventions and better understand illness course.

Conclusion and Recommendations

Depression is a highly prevalent disease that impairs all aspects of life including an individual's ability to work, relate to others, and engage in leisurely activities. Surprisingly, researchers have not developed a consensus or standard technology to

define this impairment. Furthermore, while depression-specific treatments effectively reduce impairment in psychosocial functioning, it is unclear what processes drive this change or when this change occurs in relation to improvements in depressive symptomatology across the course of illness.

We recommend that researchers and clinicians work to clarify these issues. By developing a consensus definition and measurement of psychosocial functioning, researchers and clinicians can validly and reliably assess functional impairment, improve diagnostic practices and better quantify treatment effects. Such advances foster the field's understanding of how psychosocial functioning relates to the course and treatment of depression and will increase the relevance of research findings to public health, policy decisions, and the everyday lives of individual patients.

References

American Psychiatric Association. (2000). *Diagnostic and Statistical Manual of Mental Disorders DSM-IV-TR* (Text Revision). Washington, DC: American Psychiatric Association.

Beach, S. R., & O'Leary, K. D. (1992). Treating depression in the context of marital discord: Outcome and predictors of response of marital therapy versus cognitive therapy. *Behavior Therapy, 23,* 507-528.

Bothwell, S. & Weissman, M. M. (1977). Social impairments four years after an acute depressive episode. *American Journal of Orthopsychiatry, 47*, 231-237.

De Lisio, G., Maremmani, I., Perugi, G., Cassano, G. B., Deltito, J., & Akiskal, H. S. (1986). Impairment of work and leisure in depressed outpatients. A preliminary communication. *Journal of Affective Disorders, 10,* 79-84.

Dew, M. A., Bromet, E. J., Schulberg, H. C., Parkinson, D. K., & Curtis, E. C. (1991). Factors affecting service utilization for depression in a white collar population. *Social Psychiatry and Psychiatric Epidemiology, 26,* 230-237.

Greenberg, P. E., Kessler, R. C., Birnbaum, H. G., Leong, S. A., Lowe, S. W., Berglund, P. A., & Corey-Lisle, P. K. (2003). The economic burden of depression in the United States: How did it change between 1990 and 2000? *Journal of Clinical Psychiatry, 64,* 1465-1475.

Hays, R. D., Wells, K. B., Sherbourne, C. D., Rogers, W., & Spritzer, K. (1995). Functioning and well-being outcomes of patients with depression compared with chronic general medical illnesses. *Archives of General Psychiatry, 52,*11-19.

Hirschfeld, R. M., Montgomery, S. A., Keller, M. B., Kasper, S., Schatzberg, A. F., Möller, H. J., Healy, D., Baldwin, D., Humble, M., Versiani, M., Montenegro, R., & Bourgeois, M. J. (2000). Social functioning in depression: A review. *Journal of Clinical Psychiatry, 61,* 268-275.

Imber, S. D., Pilkonis, P. A., Sotsky, S. M., Elkin, I., Watkins, J. T., Collins, J. F., Shea, M. T., Leber, W. R., & Glass, D. R. (1990). Mode-specific effects among three treatments for depression. *Journal of Consulting & Clinical Psychology, 58,* 352-359.

Kessler, R. C. (2003). Epidemiology of women and depression. *Journal of Affective Disorders, 74,* 5-13.

Miller, I. W., Keitner, G. I., Schatzberg, A. F., Klein, D. N., Thase, M. E., Rush, A. J., Markowitz, J. C., Schlager, D. S., Kornstein, S. G., Davis, S. M., Harrison, W. M., & Keller, M. B. (1998). The treatment of chronic depression, part 3: Psychosocial functioning before and after treatment with sertraline or imipramine. *Journal of Clinical Psychiatry, 59,* 608-619.

Mintz, J., Mintz, L. I., Arruda, M. J., & Hwang, S. S. (1992). Treatments of depression and the functional capacity to work. *Archives of General Psychiatry, 49,* 761-768. yes

Papakostas, G. I., Petersen, T., Mahal, Y., Mischoulon, D., Nierenberg, A. A., & Fava, M. (2004). Quality of life assessments in major depressive disorder: A review of the literature. *General Hospital Psychiatry, 26*, 13-17.

Vittengl, J. R., Clark, L. A., & Jarrett, R. B. (2004). Improvement in social-interpersonal functioning after cognitive therapy for recurrent depression. *Psychological Medicine, 34,* 643-658.

Williams, R. A., & Strasser, P. B. (1999). Depression in the workplace: Impact on employees. *American Association of Occupational Health Nurses Journal, 47,* 526-537.

Author Note

This entry was supported by Grant Number K24MH001571 (to Robin B. Jarrett, Ph.D.) from the National Institute of Mental Health. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institute of Mental Health or the National Institutes of Health.

APPENDIX B

List of Studies Excluded from the Meta-Analysis with Reasons for their Exclusion

These are studies that contained original data and assessed psychosocial functioning but were excluded from the meta-analysis. Information was provided to identify the study, describe the type of CBT used, and list the reasons for exclusion.

Study	**Type of CBT**	**Reason for Exclusion**
1. Arean et al 2005	Group CBT	Treatment effects specific to individual with MDD were not provided
2. Beach et al 1992	Ind. CBT	Treatment effects specific to individual with MDD were not provided
3. Brown et al 1984	Psychoeducation based on cognitive and behavioral elements	Treatment effects specific to individual with MDD were not provided
4. Burns et al 1994	Ind. CBT	Treatment effects specific to individual with MDD were not provided
5. Cavanagh et al 2006	Computerized CBT	Study sample was not diagnosed with MDD
6. Corney et al 2005	Ind. CBT	Study sample was not diagnosed with MDD and treatment effects specific to individuals receiving cognitive-behavioral therapy were not provided
7. Della-Posta et al 2006	Job search assistance group with CBT	Study sample was not diagnosed with MDD
8. Elkin et al 1989	Ind. CBT	Data from this study sample was excluded in order to include Imber et al 1990
9. Emanuels-Zuurveen et al 1997	Ind. CBT	Treatment effects specific to individual with MDD were not provided
10. Emanuels-Zuurveen et al 1996	Ind. CBT	Treatment effects specific to individual with MDD were not provided
11. Firth-Cozens et al 1992	Ind. CBT	Treatment effects specific to individuals receiving CBT were not provided
12. Fletcher et al 2005	Self-help manual with CBT components	Study sample was not diagnosed with MDD
13. Gallagher-Thompson et al 1990	Ind. CBT	Study did not provide acute-phase outcomes

14. Gelhart et al 2002	Group CBT	Treatment effects specific to individual with MDD were not provided
15. Gelhart et al 2001	Group CBT	Treatment effects specific to individual with MDD were not provided
16. Hagen et al 2005	Group CBT	Schizophrenia or Schizoaffective Disorder was the primary diagnosis
17. Hardy et al 1995	Ind. CBT	Study did not assess psychosocial functioning
18. Harpin et al 1982	Couples CBT	Study sample was not diagnosed with MDD
19. Kerr et al 1992	Ind. CBT	Treatment effects specific to individuals receiving CBT were not provided
20. King et al 2002	CBT techniques performed by general practitioners	Study sample was not diagnosed with MDD
21. Lenz et al 2000	Inpatient program with CBT component	Treatment effects specific to individual with MDD were not provided
22. Mataix-Cols et al 2006	Computerized CBT	Treatment effects specific to individual with MDD were not provided
23. McLean et al 1979	Ind. Behavioral Therapy	Unable to determine significance of change in psychosocial functioning from pre- to post-treatment due to insufficient data
24. Mead et al 2005	Self-help manual with CBT components	Study sample was not diagnosed with MDD
25. Miller et al 1999	Ind. CBT	Unable to determine significance of change in psychosocial functioning from pre- to post-treatment due to insufficient data
26. Miranda et al 2004	Ind. and group CBT	Treatment effects specific to individuals receiving CBT were not provided
27. Miranda et al 2003	Ind. and group CBT	Unable to determine significance of change in psychosocial functioning from pre- to post-treatment due to insufficient data
28. Mohr et al 2005	Telephone-CBT	Study sample diagnosed with Multiple Sclerosis and not with MDD

29. Moller et al 1993	Ind. CBT	Unclear if all patients received CBT and study did not assess psychosocial functioning after acute-phase treatment
30. Murphy et al 1984	Ind. CBT	Unable to determine significance of change in psychosocial functioning from pre- to post-treatment due to insufficient data
31. Nofzinger et al 1993	Ind. CBT	Data from this study was excluded in order to include data from Thase et al 1994
32. O'Leary & Beach 1990	Ind. CBT	Unable to determine significance of change in psychosocial functioning from pre- to post-treatment due to insufficient data
33. Patience et al 1995	Ind. CBT	Treatment effects specific to individuals receiving CBT were not provided and unable to determine significance of change in psychosocial functioning from pre- to post-treatment due to insufficient data
34. Proudfoot et al 2004	Computerized CBT	Study sample was not diagnosed with MDD
35. Rees et al 2005	Group CBT	Treatment effects specific to individual with MDD were not provided
36. Richards et al 2003	Self-help program based on CBT	Study sample was not diagnosed with MDD
37. Riso et al 1997	Ind. CBT	Unable to determine significance of change in psychosocial functioning from pre- to post-treatment due to insufficient data
38. Schoenbaum et al 2004	Ind. or group CBT	Unable to determine significance of change in psychosocial functioning from pre- to post-treatment due to insufficient data
39. Schoenbaum et al 2001	Ind. or group CBT	Study sample was not diagnosed with MDD, and unable to determine significance of change in psychosocial functioning from pre- to post-treatment due to insufficient data
40. Scogin et al 1989	Cognitive and behavioral bibliotherapy	Study sample was not diagnosed with MDD and did not assess psychosocial functioning
41. Scott et al 2000	Continuation CBT	No acute-phase CBT was provided

42. Simon et al 2004	Telephone CBT	Study sample was not diagnosed with MDD
43. Simon et al 1998	Behavioral activation and brief cognitive interventions	No acute phase outcomes were provided; the only outcomes provided were at 4 and 7 month follow-ups
44. Simons et al 1984	Ind. CBT	Unable to determine significance of change in psychosocial functioning from pre- to post-treatment due to insufficient data
45. Simpson et al 2000	Ind. CBT	Study sample was not diagnosed with MDD and information describing changes in psychosocial functioning for individuals receiving CBT were not provided
46. Sullivan et al 2006	Cognitive-behavioral intervention for work disability	Study sample was not diagnosed with MDD
47. Teichman et al 1995	Ind. CBT and cognitive marital therapy	Unable to determine significance of change in psychosocial functioning from pre- to post-treatment due to insufficient data
48. Thase & Simons 1992	Ind. CBT	Study did not provide acute-phase outcomes
49. Thase, Simons, McGeary et al 1992	Ind. CBT	Study did not provide acute-phase outcomes
50. Vittengl et al 2005	Ind. CBT	Unable to determine significance of change in psychosocial functioning from pre- to post-treatment due to insufficient data
51. Ward et al 2000	Ind. CBT	Study sample was not diagnosed with MDD
52. Waring et al 1990	Cognitive marital therapy	Study sample was not diagnosed with MDD
53. Yao et al 1999	CBT	Study sample was not diagnosed with MDD

References:

Arean, P. A., Ayalon, L., Hunkeler, E., Lin, E. H., Tang, L., Harpole, L., Hendrie, H., Williams, J. W., Unutzer, J., & IMPACT Investigators. (2005). Improving

depression care for older, minority patients in primary care. *Medical care, 43,* 381-390.

Beach, S. R. H., & O'Leary, K. D. (1992). Treating depression in the context of marital discord: Outcome and predictors of response of marital therapy versus cognitive therapy, *Behavior Therapy,* 23, 507-528.

Brown, R. A., & Lewinsohn, P. M. (1984). A psychoeducational approach to the treatment of depression: Comparison of group, individual, and minimal contact procedures. *Journal of Consulting and Clinical Psychology, 52,* 774-783.

Burns, D. D., Sayers, S. L., & Moras, K. (1994). Intimate relationships and depression: is there a causal connection? *Journal of Consulting and Clinical Psychology, 62,*1033-1043.

Cavanagh, K., Shapiro, D. A., Van Den Berg, S., Swain, S., Barkham, M., & Proudfoot, J. (2006). The effectiveness of computerized cognitive behavioural therapy in routine care. *British Journal of Clinical Psychology, 45,* 499-514.

Corney, R., & Simpson, S. (2005). Thirty-six month outcome data from a trial of counselling with chronically depressed patients in a general practice setting. *Psychology & Psychotherapy: Theory, Research & Practice, 78,* 127-138.

Della-Posta, C. & Drummond, P. D. (2006). Cognitive behavioural therapy increases re-employment of job seeking worker's compensation clients. *Journal of Occupational Rehabilitation, 16*, 223-230.

Elkin, I., Shea, T.M., Watkins, J.T., Imber, S.D., Sotsky, S.M., Collins, J.F., Glass, D.R., Pilkonis, P.A., Leber, W.R., Docherty, J.P., Fiester, S.J., & Parloff, M.B. (1989). National Institute of Mental Health Treatment of Depression Collaborative

Research Program: General effectiveness of treatments. *Archives of General Psychiatry, 46,* 971-982.

Emanuels-Zuurveen, L., & Emmelkamp, P. M. (1996). Individual behavioural-cognitive therapy v. marital therapy for depression in maritally distressed couples. *British Journal of Psychiatry, 169,* 181-188.

Emanuels-Zuurveen, L., & Emmelkamp, P. M. (1997). Spouse-aided therapy with depressed patients. *Behavior Modification, 21,* 62-77.

Firth-Cozens, J., & Hardy, G. E. (1992). Occupational stress, clinical treatment and changes in job perceptions. *Journal of Occupational and Organizational Psychology, 65*, 81-88.

Fletcher, J., Lovell, K., Bower, P., Campbell, M., & Dickens, C. (2005). Process and outcome of a non-guided self-help manual for anxiety and depression in primary care: A pilot study. *Behavioural and Cognitive Psychotherapy, 33,* 319-331.

Gallagher-Thompson, D. E., Hanley-Peterson, P., & Thompson, L. W. (1990). Maintenance of gains versus relapse following brief psychotherapy for depression. *Journal of Consulting and Clinical Psychology, 58,* 371-374.

Gelhart, R. P., Hand-Ronga, N., & King, H. L. (2002). Group cognitive-behavioral treatment of depression and the interaction of demographic variables. *Journal of Cognitive Psychotherapy, 16*, 469-486.

Gelhart, R. P., & King, H. L. (2001). The influence of comorbid risk factors on the effectiveness of cognitive-behavioral treatment of depression. *Cognitive and Behavioral Practice, 8,* 18-28.

Hagen, R., Nordahl, H. M., & Gråwe, R. W. (2005). Cognitive-behavioural group treatment of depression in patients with psychotic disorders. *Clinical Psychology & Psychotherapy, 12,* 465-474.

Hardy, G. E., Barkham, M., Shapiro, D. A., Reynolds, S., Rees, A., & Stiles, W. B. (1995). Credibility and outcome of cognitive-behavioural and psychodynamic-interpersonal psychotherapy. *British Journal of Clinical Psychology, 34,* 555-569.

Harpin, R. E., Liberman, R. P., Marks, I., Stern, R., & Bohannon, W. E. (1982). Cognitive-behavior therapy for chronically depressed patients. A controlled pilot study. *Journal of Nervous & Mental Disease, 170,* 295-301.

Kerr, S., Goldfried, M., Hayes, A., Castonguay, L., & Goldsamt, L. (1992). Interpersonal and intrapersonal focus in cognitive-behavioral and psychodynamic-interpersonal therapies: A preliminary analysis of the Sheffield project. *Psychotherapy Research, 2*, 266-276.

King, M., Davidson, O., Taylor, F., Haines, A., Sharp, D., & Turner, R. (2002). Effectiveness of teaching general practitioners skills in brief cognitive behavior therapy to treat patients with depression: Randomized controlled trial. *British Medical Journal, 324*, 947-950.

Lenz, G., & Demal, U. (2000). Quality of life in depression and anxiety disorders: an exploratory follow-up study after intensive inpatient cognitive behaviour therapy. *Psychopathology, 33*, 297-302.

Mataix-Cols, D., Cameron, R., Gega, L., Kenwright, M., & Marks, I. M. (2006). Effect of referral source on outcome with cognitive-behavior therapy self-help. *Comprehensive Psychiatry, 47*, 241-245.

McLean, P. D., & Hakistian, A. R. (1979). Clinical depression: comparative efficacy of outpatient treatments. *Journal of Consulting and Clinical Psychology, 47,* 818-836.

Mead, N., MacDonald, W., Bower, P., Lovell, K., Richards, D., Roberts, C., & Bucknall, A. (2005). The clinical effectiveness of guided self-help versus waiting-list control in the management of anxiety and depression: A randomized controlled trial. *Psychological Medicine, 35,* 1633-1643.

Miller, I. W., Keitner, G. I., Schatzberg, A. F., Klein, D. N., Thase, M. E., Rush, A. J., Markowitz, J. C., Schlager, D. S., Kornstein, S. G., Davis, S. M., Harrison, W. M., & Keller, M. B. (1998). The treatment of chronic depression, part 3: Psychosocial functioning before and after treatment with sertraline or imipramine. *Journal of Clinical Psychiatry, 59*, 608-619.

Miranda, J., Chung, J. Y., Green, B. L., Krupnick, J., Siddique, J., Revicki, D. A., & Belin, T. (2003). Treating depression in predominantly low-income young minority women: A randomized controlled trial. *JAMA, 290,* 57-65.

Miranda, J., Schoenbaum, M., Sherbourne, C., Duan, N., & Wells, K. (2004). Effects of primary care depression treatment on minority patients' clinical status and employment. *Archives of General Psychiatry, 61*, 827-834.

Mohr, D., Hart, S. L., Julian, L., Catledge, C., Honos-Webb, L., Vella, L., & Tasch, E. T. (2005). Telephone-administered psychotherapy for depression. *Archives of General Psychiatry, 62,* 1007-1014.

Moller, H. J., Krokenberger, M., & von Zerssen, D. (1993). Prediction of short-term outcome of neurotic-depressive inpatients: Results of an empirical study of 134 inpatients. *European Archives of Psychiatry & Clinical Neuroscience, 242*, 301-309.

Murphy, G. E., Simons, A. D., Wetzel, R. D., Lustman, P. J. (1984). Cognitive therapy and pharmacotherapy. Singly and together in the treatment of depression. *Archives of General Psychiatry, 41,* 33-41.

Nofzinger, E. A., Thase, M. E., Reynolds, C. F., Frank, E., Jennings, J. R., Garamoni, G. L., Fasiczka, A. L., & Kupfer, D. J. (1993). Sexual function in depressed men: Assessment by self-report, behavioral, and nocturnal penile tumescence measures before and after treatment with cognitive behavior therapy. *Archives of General Psychiatry, 50,* 24-30.

O'Leary, K. D., & Beach, S. R. (1990). Marital therapy: A viable treatment for depression and marital discord. *American Journal of Psychiatry, 147*, 183-186.

Patience, D. A., McGuire, R. J., Scott, A. I., & Freeman, C. P. (1995). The Edinburgh Primary Care Depression Study: Personality disorder and outcome. *British Journal of Psychiatry, 167,* 324-330.

Proudfoot, J., Ryden, C., Everitt, B., Shapiro, D. A., Goldberg, D., Mann, A., Tylee, A., Marks, I., & Gray, J. A. (2004). Clinical efficacy of computerised cognitive-

behavioural therapy for anxiety and depression in primary care: Randomised controlled trial. *British Journal of Psychiatry, 185,* 46-54.

Rees, C., McEvoy, P., & Nathan, P. R. (2005). Relationship between homework completion and outcome in cognitive behaviour therapy. *Cognitive Behaviour Therapy, 34*, 242-247.

Richards, A., Barkham, M., Cahill, J., Richards, D., Williams, C., & Heywood, P. (2003). PHASE: A randomised, controlled trial of supervised self-help cognitive behavioural therapy in primary care. *British Journal of General Practice, 53,* 764-770.

Riso, L. P., Thase, M. E., Howland, R. H., Friedman, E. S., Simons, A. D., Tu, X. M. (1997). A prospective test of criteria for response, remission, relapse, recovery, and recurrence in depressed patients treated with cognitive behavior therapy. *Journal of Affective Disorders, 43,* 131-142.

Schoenbaum, M., Unutzer, J., Sherbourne, C., Duan, N., Rubenstein, L. V., Miranda, J., Meredith, L. S., Carney, M. F., & Wells, K. (2001). Cost-effectiveness of practice-initiated quality improvement for depression: results of a randomized controlled trial. *JAMA, 286*, 1325-1330.

Schoenbaum, M., Miranda, J., Sherbourne, C., Duan, N., & Wells, K. (2004). Cost-effectiveness of Interventions for depressed Latinos. *Journal of Mental Health Policy and Economics, 7,* 69-76.

Scogin, F., Jamison, C., & Gochneaur, K. (1989). Comparative efficacy of cognitive and behavioral bibliotherapy for mildly and moderately depressed older adults. *Journal of Consulting and Clinical Psychology, 57*, 403– 407.

Scott, J., Teasdale, J. D., Paykel, E. S., Johnson, A. L., Abbott, R., Hayhurst, H., Moore, R., & Garland, A. (2000). Effects of cognitive therapy on psychological symptoms and social functioning in residual depression. *The British Journal of Psychiatry, 177*, 440-446.

Simon, G. E., Katon, W., Rutter, C., VonKorff, M., Lin, E., Robinson, P., Bush, T., Walker, E. A., Ludman, E., & Russo, J. (1998). Impact of improved depression treatment in primary care on daily functioning and disability. *Psychological Medicine, 28,* 693-701.

Simon, G. E., Ludman, E. J., Tutty, S., Operskalski, B., & Von Korff, M. (2004). Telephone psychotherapy and telephone care management for primary care patients starting antidepressant treatment: a randomized controlled trial. *JAMA, 292,* 935-942.

Simons, A. D., Levine, J. L., Lustman, P. J., & Murphy, G. E. (1984). Patient attrition in a comparative outcome study of depression: A follow-up report. *Journal of Affective Disorders, 6*, 163-173.

Simpson, S., Corney, R., Fitzgerald, P., & Beecham, J. (2000). A randomised controlled trial to evaluate the effectiveness and cost-effectiveness of counselling patients with chronic depression. *Health Technology Assessment, 4,* 1-83.

Sullivan, M. J., Adams, H., Thibault, P., Corbière, M., & Stanish, W. D. (2006). Initial depression severity and the trajectory of recovery following cognitive-behavioral intervention for work disability. *Journal of Occupational Rehabilitation, 16,* 63-74.

Teichman, Y., Bar-el, Z., Shor, H., Sirota, P., & Elizur, A. (1995). A comparison of two modalities of cognitive therapy (individual and marital) in treating depression. *Psychiatry: Interpersonal and Biological Processes, 58,* 136-148.

Thase, M. & Simons, A. (1992). Cognitive behavior therapy and relapse of nonbipolar depression: Parallels with pharmacotherapy. *Psychopharmacology Bulletin, 28*, 117-122.

Thase, M.E., Simons, A.D., McGeary, J., Cahalane, J.F., Hughes, C., Harden, T, & Friedman, E. (1992). Relapse after cognitive behavior therapy of depression: Potential implications for longer courses of treatment. *American Journal of Psychiatry, 149,* 1046-1052.

Vittengl, J. R., Clark, L. A., Jarrett, R. B. (2005). Validity of sudden gains in acute phase treatment of depression. *Journal of Consulting and Clinical Psychology, 73,* 173-182.

Ward, E., King, M., Lloyd, M., Bower, P., Sibbald, B., Farrelly, S., Gabbay, M., Tarrier, N., & Addington-Hall, J. (2000). Randomised controlled trial of non-directive counselling, cognitive-behaviour therapy, and usual general practitioner care for patients with depression. I: Clinical effectiveness. *BMJ, 321*, 1383-1388.

Waring, E. M., Carver, C., Stalker, C. A., Fry, R., & Schaefer, B. (1990). A randomized clinical trial of cognitive marital therapy. *Journal of Sex and Marital Therapy, 16,* 165-180.

Yao, D. et al. (1999). A study of depressive patients with psychological marital disorder and their treatment. *Psychological Science (China), 22,* 259-260.

APPENDIX C

Investigator's Training at the Psychosocial Research and Depression Clinic

Laboratory Membership:

- Send out Diagnostic Evaluation packets
- Assist in lab cleanup
- Learned to utilize and maintain the O:drive electronic filing system

Protection of Human Subjects in Research:

- Completed compliance training (at http://www2.utsouthwestern.edu:8081/UTSouthwesternet/)
- Completed UT Southwestern Human Subjects Tutorial (at http://cme.cancer.gov/clinicaltrials/learning/ humanparticipant-protections.asp)
- Completed HIPAA Tutorials (at http://www4.utsouthwestern.edu/hipaa/)
- Read:
    - UT Southwestern Standard of Conduct
    - The *Multiple Project Assurance of Compliance with DHHS Regulations for Protection of Human Research Subjects*
    - The book *Ethical Principles In the Conduct of Research with Human Participants*
    - The article *Ethical Principles of Psychologists and Code of Conduct*
    - *Texas Psychologist's Certification and Licensing Act*
    - The article *Data Management and Accountability in Behavioral and Biomedical Research*
    - The chapter *The Principles and Practices of Informed Consent in Patient-Oriented Research: A Primer*

Patient Care and Data Collection:

- Recruit study participants
    - Drafted and organized mailing of postcards that describe the study
    - Passed out flyers that describe the study
- Perform telephone screens with potential study participants
- Conducted 37 diagnostic evaluations
    - Trained to perform evaluations by Monica Basco, Ph.D.
    - Twenty-three were referred to alternative treatments
    - Fourteen were referred to follow-up (10 entered the study, 2 were ineligible at follow-up and referred to alternative treatments, 1 found alternative treatment before the follow-up visit, and 1 did not attend follow-up after repeated attempts at scheduling)
- One crisis intervention
- Attended weekly lab-wide clinic meetings
- Attend regularly scheduled reliability training sessions for the HRSD and SCID-I: Current MDE

Data Management and Integrity:

- Helped maintain a six-month supply of research packets for the following:
    - Pre-Diagnostic Evaluation Mail out Packets
    - Diagnostic Evaluation
    - Follow-up Diagnostic Evaluation
    - Acute Phase Cognitive Therapy Visits
    - Psychoeducation Visits

  - Blind Evaluations
  - Continuation Phase Visits
- Helped with 1st and 2nd key hand-entry and reconciliation between keyings for the following measures:
  - Beck Depression Inventory
  - Beck Hopelessness Scale
  - Inventory of Interpersonal Problems
  - List of Threatening Experiences
  - Mood Induction Visual Analog Scale
  - Spousal Criticism Scale
  - Patient Attitudes and Expectations
  - Somatic Symptoms Scale
- Perform audits of the database
- Submitted IRB modification for new recruitment strategy
- Helped write the scoring template for the RIFT
- Attend weekly lab-wide data-report meeting

Scientific Productivity and Contribution:

- Helped prepare paper presentation for the 2006 annual conference of the Society of Psychotherapy Research
- Co-authored article in *Journal of Consulting and Clinical Psychology:*

  Vittengl, J. R., Clark, L. A., Dunn, T. W., & Jarrett, R. B. (2007). Reducing relapse and recurrence in unipolar depression: A comparative meta-

analysis of cognitive therapy's effects. *Journal of Consulting and Clinical Psychology, 75,* 475-488.

- Presented poster at the 2007 UT Southwestern Graduate Student Poster Session
- Presented paper at the 2007 annual conference for the Association for Behavioral and Cognitive Therapies
- Chair an international research group focused on researching psychosocial functioning in depression
    - Group will seek to publish meta-analysis of treatment effects on psychosocial functioning in depressed populations
    - Consists of members from the University of Jena (Germany), University of Iowa, Truman State University, and The University of Texas Southwestern Medical Center at Dallas

Planned

- Administer LIFE as part of the blind evaluation
- Continue Diagnostic Evaluations
- Analyze data as described
- Papers in progress:

Dunn, T. W., Ro, E., Risch, A. K., Vittengl, J., Clark, L. A., & Jarrett, R. B. Psychosocial Functioning and Unipolar Depression: A Meta-Analysis of Changes After Treatment. To be submitted to the *Journal of Consulting and Clinical Psychology*.

Dunn, T. W., Carmody, T., & Jarrett, R. B. Changes in Psychosocial Functioning and Depressive Symptoms during Cognitive Therapy for Recurrent Depression. To be submitted to the *Journal of Clinical Psychiatry*.

Dunn, T. W. & Jarrett, R. B. Psychosocial Functioning. Chapter in *The International Encyclopedia of Depression*.

APPENDIX D

Consent Form to Participate in Diagnostic Evaluation

The University of Texas Southwestern Medical Center at Dallas

CONSENT TO PARTICIPATE IN RESEARCH

Title of Research: PROPHYLACTIC COGNITIVE THERAPY FOR DEPRESSION AND ARE COGNITIVE THERAPY'S ANTIDEPRESSANT EFFECTS DURABLE?

DIAGNOSTIC EVALUATION PHASE

Sponsor: National Institutes of Health

Investigators:	Telephone No. (Regular office hours)	Telephone No. (Other times)
1. Robin B. Jarrett, Ph.D.	214-648-5345	214-648-5555
2. Dolores Kraft, Ph.D.	469-385-7888 (all times)	
3. Greg Eaves, Ph.D.	972-596-6351 (all times)	
4. Bethany Hampton, Ph.D.	972-490-9759 (all times)	
5. Monica Basco, Ph.D.	817-781-7333 (all times)	
6. William Edwards, Ph.D.	214-361-1717	214-648-5555
7. Steven R. Krebaum, Ph.D.	214-648-0355	214-648-5555
8. Nancy Cravens, LVN	214-648-5346	214-648-5555
9. Catherine Judd, P.A.-C., M.S.	214-510-0747 (all times)	
10. Berit Johnson, Ph.D.	214-215-2499 (all times)	
11. Cindy L. Kidner, Ph.D.	214-648-5347	214-648-5555
12. Margaret Marcotte, M.S.	214-648-0176	214-648-5555
13. Todd W. Dunn, M.S.	214-648-4462	214-648-5555
14. Wendy Ringe, Ph.D.	214-648-8701	214-676-2494
15. Shelley Guess, B.A.	214-648-5343	
16. Wayne H. Denton, M.D., Ph.D.	214-648-6945	214-648-5555

INVITATION:
Medical research involves offering a plan of care to a group of patients, collecting and studying information about each patient's experience, and using that information to develop the best possible care for future patients. Because you have reported symptoms of depression, you may participate in this diagnostic evaluation. This evaluation will determine whether you are eligible to participate in a research study on Recurrent Major Depressive Disorder. The sponsor plans to include 665 patients with Recurrent Major Depressive Disorder in this study.

PURPOSE:
These pages will tell you about consenting to a diagnostic evaluation to determine whether you qualify for a research study at the Psychosocial Research and Depression Clinic of The University of Texas Southwestern Medical Center at Dallas.

You are agreeing to complete a diagnostic evaluation to determine whether you are eligible to take part in a research study on reducing depressive symptoms with cognitive therapy. In the first phase of the study, you would be treated with cognitive therapy. Cognitive therapy is a short-term talking therapy, which examines how your thoughts may influence your depressive symptoms. In the second phase of the study, if you meet study requirements, you may be randomly assigned (as by drawing straws) to: continue cognitive therapy, begin an antidepressant medication called fluoxetine (Prozac) or pill placebo, or complete follow-up evaluations every 4 months. Since you and your therapist will not select the treatment you receive in the second phase of the study, you should only participate in this evaluation if you are willing first to receive cognitive therapy for depression, and later to be randomized to receive any one of the four follow-up options: cognitive therapy, fluoxetine (Prozac), placebo, or to be followed during this time period. If any of these options are unacceptable to you, you can request a referral now and decline participating in the evaluation.

Another purpose of this evaluation is to provide you with a diagnosis of your current depressive symptoms. Afterward, clinicians can recommend appropriate treatment. There are many types of depressive illnesses and this study treats one specific depressive subtype. If you do not have this type of depression clinicians will refer you for treatment outside of this study protocol. Also, if the evaluator finds that you are not eligible to take part in research, the evaluation may be stopped and the evaluator will then provide you with a referral. It is important to understand that being ineligible for the study does not mean that your depressive symptoms do not need treatment. The clinician will encourage you to follow-up with alternative treatment if you are not eligible for our research study and help you locate such treatment.

PROCEDURES:
This evaluation will take approximately one to four hours. If you decide to participate you will be giving your permission to participate in the following procedures. This evaluation will include questions about the symptoms of depression you have now and the history of your depression. You will be asked questions about any other psychiatric symptoms you have or had in the past. It is also important to know about your present and past use of drugs and medications. The evaluation will include taking your medical history, along with your treatment history for psychiatric difficulties. We will ask you for information about your family history, and information about your current functioning in work, school, relationships, and other social activities. In addition, laboratory tests including blood and urine samples may be required which allow us to rule out some common physical causes of depression. You may be tested for pregnancy.

If you appear to be eligible for our research study, a follow-up appointment with a faculty member will be scheduled. Your evaluator may ask for you to bring a significant other (i.e., a family member or friend) to this follow-up if appropriate. This appointment will take one to three hours about a week or so after today's evaluation. The purpose of the follow-up interview is to confirm your diagnosis and talk about treatment recommendations and give you and your significant other an opportunity to ask questions

to learn more about how you can feel better. Today's evaluator will write a report based on your discussion, which will be read by the faculty evaluator. If you were asked to give blood and urine samples, your results will be reviewed by a consulting physician before this appointment. If you are not eligible, or decide not to participate in this research study at follow-up, we will provide a referral for alternate treatment.

POSSIBLE RISKS:
The risks of the evaluation process include:

1. After completing, or at anytime during the evaluation and follow-up, it may turn out that you are not eligible for this research. Being ineligible does not mean your depression is not treatable. Instead being ineligible means that this study may not or cannot provide the best treatment for the type of depression you have. The research clinicians will work with you to find you an alternate referral for your symptoms. With your written permission the results of your clinic evaluation can be mailed to the clinician who treats you.

2. If you are eligible, it takes about two weeks from the beginning of this evaluation to the time you enter treatment. If it is determined that your symptoms need more immediate attention, this evaluation will be discontinued. A referral for immediate treatment will be provided.

3. Blood samples: You may experience, discomfort, bleeding, and/or bruising. You may feel dizzy or faint. On a rare occasion, an infection may develop at the site where the blood was collected.

POSSIBLE BENEFITS:
You may benefit from this evaluation as described below.

1. You will receive a free diagnosis of your symptoms and recommendations for treatment.

2. The information you provide by completing questionnaires and providing a clinical history will help researchers learn more about the characteristics of depression.

3. You may be eligible for free treatment in a research study.

It is possible that you may not benefit from participation in this study. In the future, other people with Recurrent Major Depressive Disorder may benefit from the results of this research. New information may lead to improved medical care.

ALTERNATIVES TO PARTICIPATION IN RESEARCH:
You do not have to participate in this research study to receive care for your medical problem. Alternative care includes referral for treatment with antidepressant medication

or psychotherapy. If you decide to participate in research, but later change your mind, you will be referred for the alternative care.

THE DOCTOR'S DECISION TO STOP YOUR PARTICIPATION:
Your doctor or evaluator may stop this evaluation without your permission under any one of the follow conditions:

- Your clinician believes that participation in the study is not safe for you.
- Your clinician believes that other treatment may be more helpful.
- You fail to keep appointments and to follow the study procedures and your doctor's recommendations.

PROCEDURES AFTER STOPPING PARTICIPATION IN THE STUDY:
If you, the clinician, or the sponsor stops your participation in this research we will provide referrals for alternative treatment options.

COSTS:
The sponsor will pay the expenses for this diagnostic evaluation and laboratory work that are necessary to determine whether you are eligible for the research study. If you enter the research study, the treatment provided is free. Other expenses resulting from standard care for your medical problems are your responsibility (or the responsibility of your insurance provider or government program). There are no funds available to pay for transportation to and from the clinic, parking, lost time away from work and other activities, lost wages, or child care expenses.

PAYMENTS TO PARTICIPATE:
You will not be paid to participate in this study.

VOLUNTARY PARTICIPATION:
You have the right to agree or refuse to participate in this research. If you decide to participate and later change your mind, you are free to discontinue participation in the research at any time. Refusal to participate will involve no penalty or loss of benefits to which you are otherwise entitled. Refusal to participate will not affect your legal rights or the quality of health care that you will receive at this center.

RECORDS OF YOUR PARTICIPATION IN THIS RESEARCH:
You have the right to privacy. Any information about you that is collected for this research will remain confidential as required by law. In addition to this consent form, you will be asked to sign an "Authorization for Use and Disclosure of Protected Health Information for Research Purposes," which will contain more specific information about who is authorized to review, use, and/or receive your protected health information for the purposes of this study.

CERTIFICATE OF CONFIDENTIALITY: The principal investigator of this study, Robin B. Jarrett, Ph.D., and her associates have obtained a Certificate of Confidentiality from the Federal government. This Certificate will help researchers protect your privacy. However, the Certificate will not protect your privacy if you consent in writing to the release of information about your participation in this research to anyone else.

What is a Certificate of Confidentiality? The Department of Health & Human Services issued a Certificate of Confidentiality for this research. This Certificate enables Robin B. Jarrett, Ph.D. and/or her associates to withhold information about your participation. The protection afforded by this Certificate lasts forever. However, the Certificate will not provide protection if you consent in writing to the release of information about your participation in the research to anyone else.

Why is a Certificate of Confidentiality needed? Sensitive information about your health and the health of other members of your family may be collected and studied. The Certificate will help Robin B. Jarrett, Ph.D. and/or her associates avoid having to release identifying information about you, which could expose you and your family to unwanted financial, legal, emotional, and social consequences.

How does the Certificate of Confidentiality protect your privacy? All persons who are employed by or associated with the University of Texas Southwestern Medical Center at Dallas (and its contractors or cooperating agencies) and who have access to information about your participation in this research may withhold your name and other identifying information from all persons not connected with the conduct of that research.

This means that Robin B. Jarrett, Ph.D. and/or her associates do not have to identify you as a participant in this research in any Federal, State, or local, civil, criminal, administrative, legislative, or other proceedings.

What are the limitations of the Certificate? This Certificate does not stop you or a member of your family from identifying you as a participant in this research.

For example, if an insurance provider or employer learns about your participation in this research and obtains your consent to receive research information, Robin B. Jarrett, Ph.D. and/or her associates may not use the Certificate of Confidentiality to withhold this information.

It is important that you and your family actively protect your own privacy.

If Robin B. Jarrett, Ph.D. and/or her associates determine that you could be harmful to yourself or to others, she may report such concerns to proper authorities for your safety or the safety of others.

A Certificate of Confidentiality does not represent an endorsement of this research project by the Department of Health & Human Services or any other Federal government agency.

Could there be problems if you or someone else in the family releases information?
If you or a member of your family receives private information about you and does not maintain the privacy of that information, there is no way to predict who will have access to that private information. There is no way to predict the risks or damage which could result from unwanted release of that information.

RELEASE OF INFORMATION:
In addition to any and all authorization that you provide in the "Authorization for Use and Disclosure of Protected Health Information for Research Purposes," that you will be asked to sign as part of this research study, the confidentiality of your personally identifiable research-related information is also protected by a Certificate of Confidentiality. With this Certificate, the investigators cannot be forced (for example, by a court subpoena) to disclose information that may identify you in any federal, state, or local civil, criminal, administrative, legislative or other proceedings. However, the investigators remain free to make disclosures to protect you and others from harm.
Identifying information may be released as follows:

1. If the investigators learn about child abuse or neglect; elder abuse or neglect, or abuse or neglect of individuals in state institutions.

2. When, in our professional judgment, you may be a danger to yourself or others;

3. In the event that a suit is brought against The University of Texas Southwestern Medical Center at Dallas, which relates to your care as a research subject; or

4. In the event that you should waive your right to confidentiality by providing written consent so that you or another individual (e.g., your physician) may have access to information related to your participation in this research study.

COMPENSATION FOR INJURY:
Compensation for a physical injury resulting from participation in this research is not available from The University of Texas Southwestern Medical Center at Dallas. However, you retain your legal rights during your participation in this research.

YOUR QUESTIONS:
Robin Jarrett, Ph.D., the Principal Investigator, is available to answer your questions about this research. She can be reached at 214-648-5345. The Chairman of the IRB is available to answer questions about your rights as a participant in research or to answer your questions about an injury or other complication resulting from your participation in this research. You may telephone the Chairman of the IRB during regular office hours at 214-648-3060.

YOU WILL HAVE A COPY OF THIS SIGNED AND DATED CONSENT FORM TO KEEP.

Your signature indicates that you have read (or been read) the information provided above, that you have received answers to all of your questions, and that you have freely decided to participate in this research. By agreeing to participate in this research, you are not giving up any of your legal rights.

Again, in signing this consent you are also agreeing with the following statement:
If the research team cannot locate me, I give my permission for the team to contact the people listed below in order to locate me.

Printed Name and Signature of Subject Age Date

Subject Address Subject Telephone Number

Email Address

Printed Name and Signature of Person Obtaining Consent/Investigator Date

Contact #1 (Name / Relation) Address Telephone

Contact #2 (Name / Relation) Address Telephone

Investigator's Statement:
I certify to the best of my knowledge that the information provided is accurate and up to date. I have examined _____, and it is my judgment that he/she is competent to give informed consent to participate in this study.

Signature of Investigator Date

APPENDIX E

Consent Form to Participate in Randomized Controlled Trial

The University of Texas Southwestern Medical Center at Dallas

CONSENT TO PARTICIPATE IN RESEARCH

Title of Research: PROPHYLACTIC COGNITIVE THERAPY FOR DEPRESSION AND
ARE COGNITIVE THERAPY'S ANTIDEPRESSANT EFFECTS DURABLE?

Sponsor: National Institutes of Health

Investigators:	Telephone No. (Regular office hours)	Telephone No. (Other times)
1. Robin B. Jarrett, Ph.D.	214-648-5345	214-648-5555
2. Dolores Kraft, Ph.D.	469-385-7888 (all times)	
3. Greg Eaves, Ph.D.	972-596-6351 (all times)	
4. Bethany Hampton, Ph.D.	972-490-9759 (all times)	
5. Monica Basco, Ph.D.	817-781-7333 (all times)	
6. William Edwards, Ph.D	214-361-1717	214-648-5555
7. Steven R. Krebaum, Ph.D.	214-648-0355	214-648-5555
8. Nancy Cravens, L.V.N.	214-648-5346	214-648-5555
9. Catherine Judd, P.A.-C., M.S.	214-510-0747 (all times)	
10. Berit Johnson, Ph.D.	214-215-2499 (all times)	
11. Cindy L. Kidner, Ph.D.	214-648-5347	214-648-5555
12. Margaret Marcotte, M.S.	214-648-0176	214-648-5555
13. Todd W. Dunn, M.S.	214-648-4462	214-648-5555
14. Wendy Ringe, Ph.D.	214-648-8701	214-676-2494
15. Shelley Guess, B.A.	214-648-5343	
16. Wayne H. Denton, M.D., Ph.D.	214-648-6945	214-648-5555

INVITATION:
Medical research involves offering a plan of care to a group of patients, collecting and studying information about each patient's experience, and using that information to develop the best possible care for future patients. Because you have recurrent major depressive disorder, you may be eligible to participate in this research study. The research study evaluates how effective 16-20 sessions of cognitive therapy, with or without additional treatment, is in reducing the chance that your depressive symptoms return. The sponsor plans to include 665 patients with recurrent major depressive disorder in this study.

PURPOSE:
Cognitive therapy is a short-term talking therapy, which examines how your thoughts may be influencing your depressed mood. This treatment is a form of psychotherapy in which you set treatment goals and complete homework assignments. The goal of therapy

is to reduce your depressive symptoms and to teach you skills which may decrease the chance of getting depressed again in the future.

You were selected as a possible participant because your diagnosis is recurrent major depressive disorder, meaning you have had at least two episodes of depression during your lifetime. One of the risks of recurrent major depressive disorder is that sometimes depressive symptoms return even after a person's mood has been normal. When depressive symptoms return, it is called a relapse or a recurrence, depending on how long the individual has been free from depressive symptoms.

Previous work in research studies has shown that 16-20 sessions of cognitive therapy can reduce depressive symptoms completely in about 65% of patients who suffer from recurrent depression; other depressions may improve partially. For some patients, cognitive therapy can be as effective as antidepressant medication in reducing symptoms. More research is needed to confirm the effectiveness and amount needed of cognitive therapy for reducing the chance of relapse/recurrence of depression.

The purpose of this study is to determine how to best reduce relapse/recurrence for each individual. The "test" in the study involves comparing which treatments are most effective for each individual. After completing 16-20 sessions of cognitive therapy, if eligible, you will be randomly assigned (as by drawing straws) to one of the following options for the next 8 months: (a) 10 additional "booster" sessions of cognitive therapy; (b) 10 sessions in which antidepressant medication is prescribed; (c) 10 sessions in which a placebo (an inactive substance) is prescribed; or (d) you will not be randomized and instead you will have follow-up evaluations every 4 months. This 8-month experimental period is followed by a 24-month follow-up phase during which you may receive diagnostic evaluations every 4 months. The duration of your follow-up will be determined by the investigators.

PROCEDURES:
What does my participation involve? A Significant Time Commitment to Complete Study Procedures

Your participation in the study could last as long as 36 months and as few as 3 months. In any case, participating in the study requires a large commitment of your time. You should only agree to participate if you are able to invest an average of 3 hours a week over the next 36 months to complete questionnaires and attend diagnostic and treatment sessions during standard business hours. During some phases of the study you will spend more or less than 3 hours a week completing study procedures.

Participation in this study is divided into three phases. Everyone will participate in the first phase (16-20 sessions of cognitive therapy). If eligible, you will be randomized to one of four conditions over a period of 8 months, and finally you will complete a 24-month follow-up phase. Since you and your therapist will not select the treatment option

to which you will be assigned during the second phase, you should only participate if any of the four treatment options would be acceptable to you.

Two of the treatment strategies involve more treatment and two of the strategies do not. The purpose of this study is to determine which patients need more or less treatment after 16-20 sessions of cognitive therapy.

If you complete all three phases, you would participate for approximately 36 months. To participate in any phase of this study, you must be evaluated and judged to be eligible according to study criteria. Either you or the investigators can decide at any point that you need to withdraw from the study. At any point you can request a referral for alternative treatment. Should you experience a relapse or recurrence during Phase 2 or 3 you will continue to be followed by research staff and aided in locating appropriate treatment and you will be referred for treatment outside the study.

Maintaining Contact with Researchers for as Long as 36 Months

Finally, if you agree to participate in this study, you will also be agreeing to maintain contact with the research team for as long as 36 months. We will ask you to provide the names, addresses, and telephone numbers of two people who would always know your address and telephone number, in the event that the research team has difficulty locating you. When speaking with the contact person, we would identify ourselves as "researchers at The University of Texas Southwestern Medical Center at Dallas" but would not disclose any additional details regarding your participation in this study. If at any point you wish to withdraw from the study and receive no more contact from the research team, you may.

Agreeing to Fulfill Study Requirements

All treatment and evaluation in this research protocol are free of charge, in exchange for completing all study procedures on time. Failure to complete study procedures on time will result in being withdrawn from the study and referred for treatment outside the study. The following summarizes the conditions of participation, or what is expected of you during the study.

1. You must be able to attend therapy and clinic appointments during business hours (8 A.M. - 5 P.M., Monday-Friday).

2. If you need to cancel a therapy or clinic appointment, you agree to notify your therapist or evaluator at least 24 hours in advance. The therapists who provide the treatment in the study have full clinical practices and need to have at least a 24-hour advance notice if you need to cancel a session. Frequently, the clinicians have a "waiting list" for patients wanting to receive treatment. The 24 hour notice is required so that your clinician can allow another patient to reschedule in

your place. Failure to provide a 24-hour notice for cancellation may cause you to be withdrawn from the study.

3. You are agreeing to complete study questionnaires throughout the duration of the study. You also agree to answer questionnaires completely and turn them in on time. Failure to do this may cause you to be withdrawn from the study.

4. To the best of your knowledge, you will be living in the Dallas area for at least 36 months so that you can continue to attend clinic evaluations.

5. You agree to be randomized to either booster sessions of cognitive therapy, fluoxetine, placebo, or evaluation only if you are eligible at the end of 16-20 sessions of cognitive therapy in Phase 1.

6. You agree to utilize medically approved birth control methods (e.g., birth control pills, IUD, sterilization, diaphragm, or condom with spermicide) and will notify clinic staff immediately if you should become pregnant. After 16-20 sessions of cognitive therapy and prior to randomization into the second phase of the study, all eligible women of child bearing potential will undergo a urine pregnancy test in order to avoid assigning a pregnant woman to medication treatment, which could be harmful to the fetus or developing child. Women who are breast feeding or pregnant, are not eligible to participate in this study. If you become pregnant during the study, you may be exposing yourself or your child to unknown risks if you are assigned to take medication. You must notify your physician immediately if you: (a) become pregnant; (b) think you may have become pregnant; (c) plan to become pregnant; or (d) plan to discontinue contraception. Should you become pregnant during the study, you may need to stop taking the study medication. You should discuss this with your study psychiatrist immediately. If you decide to discontinue the medication, you may continue in the study for follow-up.

7. You agree to avoid using mood altering drugs such as cocaine, cannabis, opiates, amphetamines, and barbiturates. In addition you agree to avoid using alcohol and to only use over-the-counter products or prescription medication, which have been approved by the study physician.

8. By participating in this research study, you are agreeing to forego any treatment for psychiatric difficulties (both therapy and medication) which is not provided within this study prior to discussing this with the Project Coordinator and your clinician.

The following pages provide more detail about each of the phases which have been described above. As you progress through each phase, research staff will review what your participation entails before you enter the next phase.

PHASE 1: Purpose -- Symptom Reduction; Method – 16-20 Sessions of Cognitive Therapy

In this first phase you will be asked to see your therapist twice per week for cognitive therapy. The therapy will be reduced to once per week in the latter part of your treatment. You will be treated for 16-20 sessions. Your therapist will consult with the investigators and let you know when your treatment frequency will be reduced to once per week, and how many sessions (between 16 and 20) you will receive. Your therapy will be delivered by experienced therapists. The purpose of Phase 1 will be to evaluate which types of depression respond completely to cognitive therapy. The sessions are designed to reduce the severity of depressive symptoms and to teach you skills, which may decrease the chance that you experience major depressive disorder in the future. Each session will last about 60 minutes and will be videotaped for research purposes. There are no expected discomforts or inconveniences other than the need to participate in treatment, to attempt the homework assignments, which are designed by you and your therapist, and to complete a packet of questionnaires. You deliver the questionnaires to your therapist at the beginning of the session in exchange for your treatment. If you miss a session, you will be asked to reschedule it. Therapy is likely to be more successful if you stay "on schedule" and complete your homework.

There will be no fee for cognitive therapy sessions. However, your time will be required for completing the packet of questionnaires prior to the session. While we cannot and do not guarantee or promise that you will respond to this treatment, we would like for you to try this treatment for all 16-20 sessions, or approximately 12 weeks. If at the end of treatment your symptoms have not improved enough or you are dissatisfied with the treatment, we will refer you for alternate treatment.

If you are unable to complete 16-20 sessions of cognitive therapy then you will be referred for appropriate treatment outside of the study. If you are ineligible for Phase 2 (for example, because it is determined that you would benefit from additional treatment) you will be referred for appropriate treatment outside the study.

After approximately 7-8 weeks of treatment, you will be scheduled for a visit in this clinic with the evaluators you met when you entered the study. The purpose of this visit is to “check-in” with you about your progress in the study and to review what will happen to you after the 16-20 sessions of cognitive therapy ends. You will also complete study forms during this visit which should take approximately 1-2 hours.

PHASE 2: Purpose--Test of Symptom Protection; Method--Randomization to One of Four Treatment Strategies: 8 Months

If you are able to complete 16-20 sessions of cognitive therapy and are eligible to continue in the study, Phase 2 will begin. You will be randomly assigned (as by drawing straws) to receive either 8 months of: (a) cognitive therapy booster sessions; (b) an antidepressant medication called fluoxetine (Prozac) plus clinical management; (c)

placebo (an inactive substance) plus clinical management; or (d) evaluations only every 4 months. Neither you, nor your therapist, nor any of your evaluators will know at the beginning of the study to which treatment strategy you may be assigned. In all treatment strategies, you will be asked to complete questionnaires and answer questions about your symptoms and life.

If it is necessary for you to be treated outside of this study, the research team will request your written permission to release and obtain clinical information from the practitioner providing your treatment. If it is necessary for you to receive treatment outside of this protocol, we will ask you to complete all phases of the study, and continue to have your symptoms and progress monitored by research staff.

Booster Sessions of Cognitive Therapy

One of the purposes of this study is to determine which patients require additional booster sessions of cognitive therapy to prevent early relapse or recurrence, and how this type of treatment compares with alternate treatment (e.g., medication) or no additional treatment. If you are assigned to booster sessions, you will be asked to participate in therapy every other week for 2 months, followed by 6 months of monthly sessions. If you find that your depressive symptoms are returning between sessions, you may contact your therapist to schedule one "emergency" session during the course of 8 months. In all you will have 10 booster sessions in 8 months. Sessions will be scheduled for approximately 90 minutes. If you experience a relapse or recurrence during this phase of the study, additional treatment may be necessary and you will complete any remaining booster sessions.

If you are eligible for additional treatment within the study, that treatment would be described in a separate consent form and you would receive the treatment only if you are eligible and if you agreed to do so. If you need treatment outside of the study, you will be referred and research staff will continue to monitor your symptoms for the duration of the study.

Fluoxetine (Prozac) or Placebo Plus Clinical Management

An additional purpose of this study is to determine which patients would benefit from additional treatment with fluoxetine (Prozac) following 16-20 initial sessions of cognitive therapy. The question is whether such treatment with fluoxetine prevents early relapse or recurrence. Neither you nor your psychiatrist nor the evaluators will know whether you are taking fluoxetine or a placebo. Placebo tablets will be identical in appearance to fluoxetine tablets. You will be seen for ten sessions by a psychiatrist who will prescribe fluoxetine or a placebo. You will be seen for an initial session lasting 30-45 minutes, and nine additional sessions lasting 15-30 minutes. The frequency of these sessions will be two times per month for two months and once per month for six additional months for a total of 10 sessions. You will initially receive 10 mg/day of fluoxetine. This medication is the most widely used modern antidepressant and has been shown to be beneficial in

preventing early relapse or recurrence. Your dose of fluoxetine will be adjusted by your response to the medication and any side effects you experience from a minimum of 10 mg/day to a maximum of 40 mg/day. If you are assigned to the placebo group, the same methods will be followed but you will not be taking an antidepressant medication.

At each session the treating psychiatrist will question you about your symptoms and any possible medication side effects and your blood pressure, pulses, and weight will be monitored. Please see "Possible Risks" on page 8 of this consent form for a summary of potential side effects related to fluoxetine.

You will be asked to take your medication every day as prescribed. You should not take any other medications, either over the counter or prescribed by another physician, without the knowledge of the study psychiatrist. If other non-study medications are taken, they could cause side effects. In addition, you should refrain from using alcohol while in the study. Your medication should be stored safely outside of the reach of children.

If you should have a relapse or recurrence during this phase, when necessary the blind will be broken to a non-study clinician (i.e. he/she will be told whether you received placebo or fluoxetine) so that appropriate treatment decisions can be made for you. You would continue to be followed for the duration of the study by research staff.

Follow-Up Only (Blind Evaluations)

If you are assigned to evaluation without further treatment, you will be asked to come to the clinic every four months for an evaluation (see blind evaluation below). You will have two evaluation sessions in eight months. These sessions last approximately 2-3 hours. If during these 8 months it is necessary to begin treatment for depression or any other disorders, research staff can and will assist you in finding appropriate treatment.

PHASE 3: Purpose -- Symptom Detection; Method -- Long-term Follow-up

Phase 3 of the study involves diagnostic blinded evaluations every 4 months for 24 months (see below). The research purpose of the long-term follow-up program is to determine how long you can stay well, symptom free. Your participation in the long-term follow-up program will last 24 months. The Project Coordinator will help you stay on schedule. During this time you will come in to the clinic, you will complete a packet of questionnaires (before you come to the clinic) and your symptoms will be evaluated by a trained evaluator.

Periodic "Blind" Clinical Evaluation:

All subjects will have a "blind" clinic evaluation when Phase 1 is completed, twice during Phase 2 of the study (4 and 8 months after randomization), and six times during Phase 3 (12, 16, 20, 24, 28 and 32 months after randomization). This means that you will receive clinical evaluations conducted by an evaluator who does not know anything about

your treatment history. Please do not tell the evaluator anything about the type of treatment you have received and do not mention your doctor's or evaluator's name. The purpose of this interview is to evaluate your condition and symptoms without any bias or prejudice about your treatment.

At your first blind evaluation, you will be asked to listen to music, to remember a sad time during your life, and to complete some questionnaires. If you consent to this mood induction procedure, you will complete it at the first, third and last blind evaluations. All blind evaluations will be videotaped.

If you relapse or recur during phase 2 or 3 of the study, you will be referred for treatment outside the study or, if eligible, you will be offered treatment within the protocol and a separate consent form will describe the treatment. In either case, you will continue to be followed for the entire study period.

POSSIBLE RISKS:
The following outlines the risks of this research study for each phase.

Phase 1: The risk of Phase 1 is that you may not respond to cognitive therapy. As stated previously, approximately 65% of patients with unipolar depression respond to cognitive therapy. If you do not respond during the acute phase of treatment, we will provide referrals.

Phase 2:
1. Relapse/Recurrence: A risk of Phase 2 is that you could have a relapse or recurrence of your depressive symptoms. If this should occur, we will refer you for additional treatment outside of this study, or if you are eligible, we will offer you another protocol, while we continue to monitor your symptoms.

2. Fluoxetine: Fluoxetine, a Food and Drug Administration (FDA) approved antidepressant medication for the treatment of major depression, is now one of the most widely prescribed antidepressants in the United States. Side effects, which might occur when taking this medication, are: dry mouth, nervousness, headache, nausea, drowsiness, weight loss, diarrhea, constipation, and sexual dysfunction. Symptoms, which may also occur briefly if the medication is discontinued, include insomnia, fatigue, abdominal cramps, increased anxiety, a flu-like malaise, diarrhea, and recurrence of depressed mood. Placebo and fluoxetine may have similar side effects. Although fluoxetine is generally a safe medication, extremely rare reactions may include liver inflammation, seizures, or reduced blood cell counts. These risks do not differ from those faced by patients taking fluoxetine in non-research settings. In addition, antidepressant medications have been shown to impair mental and physical abilities required for the performance of potentially hazardous tasks such as operating machinery or driving an automobile. Although this is not typically associated with fluoxetine, caution is encouraged.

The FDA has warned that some patients taking antidepressant medication have been observed to show worsening of depressive symptoms and suicidal impulses. Additional symptoms observed have included: anxiety, agitation, panic attacks, insomnia, irritability, hostility, aggressiveness, impulsivity, restlessness, and euphoria. While the FDA has not established a causal link between these symptoms and either the worsening of depression or suicidal impulses, the FDA warns that these symptoms may be more likely when fluoxetine is either started, the dose is increased, or discontinued. Careful monitoring is important, especially at these times.

The FDA recommends that all patients taking antidepressants be monitored not only by their treating physicians but also by family members or close associates. If you, your family member, or your close associate notice any of the symptoms listed above, suicidal ideas or impulses, or other changes in behavior, call the study psychiatrist or other research personnel immediately. If you need immediate assistance, go to the nearest emergency room.

The risks of treatment are balanced in part by a careful assessment of your health and ongoing examination of your response to medication. You will be informed of any changes in the way the study is conducted and of any new risks to which you could be exposed. It is possible that your depression may not improve with fluoxetine. In that case, we will recommend alternative treatment.

3. Placebo: If you receive a placebo, you will not receive active medication for your condition. If your condition becomes worse, your participation in the study may stop at the discretion of the investigators. If this happens, your doctor can discuss alternative care with you.

4. Blood samples: You may experience, discomfort, bleeding, and/or bruising. You may feel dizzy or faint. On a rare occasion, an infection may develop at the site where the blood was collected.

5. Unforeseen risks: A previously unknown side effect may occur. A side effect because of an interaction of fluoxetine with other medications you take (prescribed or over-the-counter) may result from your participation in the study. It is not possible to estimate the chances of such occurrences or their severity.

6. Risks to an unborn child or a breast-fed infant: It is not known whether fluoxetine may harm an embryo or fetus or an infant who is breast-feeding. It is not known whether treatment with fluoxetine may lead to birth defects.

7. Women: A woman who is pregnant or is breast-feeding an infant may not participate in this research. A pregnancy test will be performed for any woman who is capable of bearing a child and wishes to participate in this study. If you are a woman who can bear children and suspect pregnancy during the time you receive treatment in this study, please tell your doctor immediately.

Pregnancy During the Acute Phase

If you become pregnant during the acute phase of treatment, you will finish the acute treatment but will be ineligible for randomization into Phase 2 of the research study. This avoids your exposure to fluoxetine during pregnancy. Instead you will receive 10 booster sessions of cognitive therapy over 8 months. The frequency of these sessions will be two times per month for two months and once per month for six additional months for a total of 10 sessions. Each session will last 60-90 minutes. Eight months after your initial randomization, you will enter longitudinal follow-up where you will be seen for "blind" evaluations every 4 months for 24 months.

Pregnancy During the Continuation Phase

Cognitive Therapy and Follow-Up Only Conditions

If you become pregnant during the 8-month continuation phase of these conditions you will continue with your treatment as planned.

Fluoxetine and Pill Placebo Conditions

If you become pregnant during the 8-month continuation phase of cognitive therapy or follow-up, you will be seen by a second psychiatrist who has not been associated with your treatment in this protocol. The blind will be broken to this physician so you will know whether the fetus has been exposed to fluoxetine. You (and your partner, if applicable) and the psychiatrist will review your situation and make a decision about whether you should continue fluoxetine (as applicable) or switch to another treatment (e.g., booster sessions of cognitive therapy). If you have been taking pill placebo, you may also be switched to booster sessions of cognitive therapy. If you are treated with booster sessions of cognitive therapy, you will receive booster sessions of cognitive therapy until you complete 8 months of treatment after being randomized. Each session will last 60-90 minutes. Eight months after your initial randomization, you will enter longitudinal follow-up and you will continue to come to the clinic every 4 months for a "blind" evaluation.

If you decide to continue taking fluoxetine, you understand that the potential risks to the fetus are not yet known.

Pregnancy during Longitudinal Follow-Up (24 months)

If you become (or plan to become) pregnant during this 24-month phase you will continue in the study as planned.

- **Avoiding pregnancy:** Please ask your doctor or other research personnel about the appropriate ways to avoid becoming pregnant during participation in this study.

Phase 3: The risk of Phase 3 is that you could have a relapse or recurrence of your depressive symptoms. If this should occur, we may offer you alternative treatment or provide referrals.

By giving the research team the name and address of two people who always know where to find you, there is a slight risk that these persons could associate you with the Psychosocial Research and Depression Clinic. We will guard against this by only identifying ourselves as "researchers at The University of Texas Southwestern Medical Center at Dallas".

How you can help reduce some of the risks: During your participation in this study, research personnel will watch closely to determine whether there are complications that need medical care. It is your responsibility to do the following:

- Ask questions about anything you do not understand.
- Keep appointments.
- Follow the clinician's recommendations.
- Let us know if your telephone number changes.
- Store study pills in a secure place at home away from anyone who is unable to read and understand labels, especially children.
- Tell your study doctor before taking any new medication even if the medication is prescribed by another doctor for a different medical problem.
    - Tell your regular doctor about your participation in this study.
- Educate yourself, your family member, and/or close associate about: a) the risk of depression worsening or developing suicidal thoughts or impulses and b) how to contract research personnel or the nearest emergency room should these symptoms occur.

What to do if you have problems: If you experience unusual symptoms, pain or any other problems at any time during your participation in the study, your doctor can recommend treatment. Please report the problem to your doctor promptly. *Telephone numbers where study clinicians may be reached are listed above.*

POSSIBLE BENEFITS:
The benefit you will receive during Phase 1 of the study is cognitive therapy provided in exchange for completing study evaluations. If you do not respond to cognitive therapy, you will be referred for appropriate treatment.

The benefit you will receive during Phase 2 of the study is additional treatment and/or intensive monitoring of your symptoms, with referral for appropriate treatment should you relapse or symptoms recur.

The benefit you will receive during Phase 3 of the study is intensive monitoring of your symptoms.

You will be paid $20 for the blind evaluations you complete. You will be paid an additional $10 at the first, third, and last blind evaluations if you participate in the mood induction procedure.

Your participation will benefit society by allowing researchers to compare viable, alternative treatment strategies to reduce the chance of depressive relapse/recurrence.

ALTERNATIVES TO PARTICIPATION IN RESEARCH:
Phase 1 -- to reduce depressive symptoms. Besides cognitive therapy there are alternative treatments which effectively reduce depressive symptoms. These include medications, some longer-term psychotherapies, and the combination of medications plus psychotherapy. If you agree to participate in the study, then you will be agreeing to discontinue and postpone alternative treatments.

Phase 2 -- to prevent new depressive symptoms. Pharmacotherapists in clinical practice routinely recommend that patients with recurrent depression receive antidepressant medication for 6-12 months after they feel normal to prevent depressive symptoms from returning. Cognitive therapists in clinical practice recommend that patients with recurrent depression return for periodic booster sessions and/or return for additional sessions at the first sign of depressive symptoms.

Phase 3 -- to detect depressive episodes. The monitoring of syndromes and symptoms in this study is more intensive than procedures used in clinical practice. If depression returns, early intervention can increase treatment effectiveness.

If you prefer any of the alternatives to study procedures used during any phase, you should request a referral and forego participating in this study.

THE DOCTOR'S DECISION TO STOP YOUR PARTICIPATION:
Your doctor or the sponsor may stop your participation in this research without your permission under any one of the follow conditions:

- Your medical problems remain unchanged or become worse.
- Side effects become very severe.
- Your clinician believes that participation in the study is not safe for you.
- Your clinician believes that other treatment may be more helpful.

- The sponsor or the Food and Drug Administration (FDA) stops the research for safety.
- The sponsor cancels the study.
- You fail to keep appointments and to follow the study procedures and your doctor's recommendations.

PROCEDURES AFTER STOPPING PARTICIPATION IN THE STUDY:
If you, the clinician, or the sponsor stops your participation in this research, it is your responsibility to come to the clinic for final evaluation and discussion about future treatment. At that time, please return any unused study medication, including empty containers.

COSTS:
The sponsor will pay the expenses for treatment and evaluation that are part of this study. Other expenses resulting from standard care for your medical problems are your responsibility (or the responsibility of your insurance provider or government program). There are no funds available to pay for parking expenses, transportation to and from the clinic, lost time away from work and other activities, lost wages, or child care expenses.

PAYMENTS TO PARTICIPATE:
If you are randomized after acute treatment, you will be paid $20 for each blind evaluation. You will not be paid for any other visits during the study.

UT Southwestern, as a State agency, will not be able to make any payments to you for your participation in this study if the State Comptroller has issued a "hold" on all State payments to you. Such a "hold" could result from your failure to pay child support payments, student loans, franchise taxes, etc. Should this occur, UT Southwestern will be able to pay you for your participation in this study after you have made the outstanding payments, and the State Comptroller has issued a release of the "hold."

VOLUNTARY PARTICIPATION:
You have the right to agree or refuse to participate in this research. If you decide to participate and later change your mind, you are free to discontinue participation in the research at any time. Refusal to participate will involve no penalty or loss of benefits to which you are otherwise entitled. Refusal to participate will not affect your legal rights or the quality of health care that you will receive at this center. Any significant new information, which becomes available during your participation in the research and may affect your health, safety, or willingness to continue in the study will be given to you.

RECORDS OF YOUR PARTICIPATION IN THIS RESEARCH:
You have the right to privacy. Any information about you that is collected for this research will remain confidential as required by law. In addition to this consent form, you will be asked to sign an "Authorization for Use and Disclosure of Protected Health Information for Research Purposes," which will contain more specific information about

who is authorized to review, use, and/or receive your protected health information for the purposes of this study.

CERTIFICATE OF CONFIDENTIALITY: The principal investigator of this study, Robin B. Jarrett, Ph.D., and her associates have obtained a Certificate of Confidentiality from the Federal government. This Certificate will help researchers protect your privacy. However, the Certificate will not protect your privacy if you consent in writing to the release of information about your participation in this research to anyone else.

What is a Certificate of Confidentiality? The Department of Health & Human Services issued a Certificate of Confidentiality for this research. This Certificate enables Robin B. Jarrett, Ph.D. and/or her associates to withhold information about your participation. The protection afforded by this Certificate lasts forever. However, the Certificate will not provide protection if you consent in writing to the release of information about your participation in the research to anyone else.

Why is a Certificate of Confidentiality needed? Sensitive information about your health and the health of other members of your family may be collected and studied. The Certificate will help Robin B. Jarrett, Ph.D. and/or her associates avoid having to release identifying information about you which could expose you and your family to unwanted financial, legal, emotional, and social consequences.

How does the Certificate of Confidentiality protect your privacy? All persons who are employed by or associated with the University of Texas Southwestern Medical Center at Dallas (and its contractors or cooperating agencies) and who have access to information about your participation in this research may withhold your name and other identifying information from all persons not connected with the conduct of that research.

This means that Robin B. Jarrett, Ph.D. and/or her associates do not have to identify you as a participant in this research in any Federal, State, or local, civil, criminal, administrative, legislative, or other proceedings.

What are the limitations of the Certificate? This Certificate does not stop you or a member of your family from identifying you as a participant in this research.

For example, if an insurance provider or employer learns about your participation in this research and obtains your consent to receive research information, Robin B. Jarrett, Ph.D. and/or her associates may not use the Certificate of Confidentiality to withhold this information.

It is important that you and your family actively protect your own privacy.

If Robin B. Jarrett, Ph.D. and/or her associates determine that you could be harmful to yourself or to others, she may report such concerns to proper authorities for your safety or the safety of others.

A Certificate of Confidentiality does not represent an endorsement of this research project by the Department of Health & Human Services or any other Federal government agency.

Could there be problems if you or someone else in the family releases information?
If you or a member of your family receives private information about you and does not maintain the privacy of that information, there is no way to predict who will have access to that private information. There is no way to predict the risks or damage which could result from unwanted release of that information.

RELEASING INFORMATION:
In addition to any and all authorization that you provide in the "Authorization for Use and Disclosure of Protected Health Information for Research Purposes," that you will be asked to sign as part of this research study, the confidentiality of your personally identifiable research-related information is also protected by a Certificate of Confidentiality. With this Certificate, the investigators cannot be forced (for example, by a court subpoena) to disclose information that may identify you in any federal, state, or local civil, criminal, administrative, legislative or other proceedings. However, the investigators remain free to make disclosures to protect you and others from harm. Identifying information may be released as follows:

1. If the investigators learn about child abuse or neglect; elder abuse or neglect, or abuse or neglect of individuals in state institutions.
2. When, in our professional judgment, you may be a danger to yourself or others;
3. In the event that a suit is brought against The University of Texas Southwestern Medical Center at Dallas, which relates to your care as a research subject; or
4. In the event that you should waive your right to confidentiality by providing written consent so that you or another individual (e.g., your physician) may have access to information related to your participation in this research study.

AUDIO AND VIDEOTAPES:
Therapy sessions and clinic or telephone evaluations will be audiotaped or videotaped. These audiotapes, videotapes, and questionnaires will be used for research, teaching, and medical purposes only. Each tape will be labeled with your study identification number and not your name and your name will not be disclosed to anyone other than study personnel.

COMPENSATION FOR INJURY:
Compensation for a physical injury resulting from participation in this research is not available from the University of Texas Southwestern Medical Center at Dallas. However, you retain your legal rights during your participation in this research.

YOUR QUESTIONS:
Robin Jarrett, Ph.D. is available to answer your questions about this research. The Chairman of the IRB is available to answer questions about your rights as a participant in research or to answer your questions about an injury or other complication resulting from your participation in this research. You may telephone the Chairman of the IRB during regular office hours at 214-648-3060.

YOU WILL HAVE A COPY OF THIS SIGNED AND DATED CONSENT FORM TO KEEP.

Your signature indicates that you have read (or been read) the information provided above, that you have received answers to all of your questions, and that you have freely decided to participate in this research. By agreeing to participate in this research, you are not giving up any of your legal rights.

Again, in signing this consent you are also agreeing with the following statements:
If the research team cannot locate me, I give my permission for the team to contact the people listed below in order to locate me. Furthermore, I agree to the recording of my evaluation or therapy sessions and the use of such audiotape or videotape for medical, educational, and research purposes.

Printed Name and Signature of Subject Age Date

Subject Address Subject Telephone Number

Email Address

Printed Name and Signature of Person Obtaining Consent/Investigator Date

Contact #1 (Name / Relation) Address Telephone

Contact #2 (Name / Relation) Address Telephone

Investigator's Statement: I certify to the best of my knowledge that the information provided is accurate and up to date. I have examined _____, and it is my judgment that he/she is competent to give informed consent to participate in this study.

Signature of Investigator Date

APPENDIX F

The Longitudinal Interval Follow-up Evaluation – Range of Impaired Functioning Tool and Scoring Template

The Longitudinal Interval Follow-up Evaluation – Range of Impaired Functioning Tool (RIFT)

(1*a*) Employment: _____
Which of the following categories best characterizes the degree to which the patient's current (past week) work activities have been impaired as a result of psychopathology?
0 = Not applicable. Did not work during the past week, for reasons other than psychopathology.
1= No impairment – high level. Worked as much as someone in his social situation would be expected to work, and worked at a high level.
2 = No impairment – satisfactory level. Worked as much as someone in his social situation would be expected to work, and worked at a satisfactory level.
3 = Mild impairment. Worked somewhat less than someone in his social situation would be expected to work and/or had mild difficulties in carrying out work activities.
4 = Moderate impairment. Has missed a lot of work and/or has had considerable difficulties in carrying out work activities.
5 = Severe impairment. Has missed a great deal of work when someone in his social situation would have been expected to work and/or has been virtually unable to carry out his work activities when he did work.
6 = No information.

(1*b*) Household: _____
Which of the following categories best characterizes the degree to which the patient's current (past week) household work has been impaired as a result of psychopathology?
0 = Not applicable. Did not carry out household duties during the past week for reasons other than psychopathology.
1 = No impairment – high level. Has carried out housework most of the time that would be expected, and worked at a high level.
2 = No impairment – satisfactory level. Has carried out housework most of the time that would be expected, and worked at a satisfactory level.
3 = Mild impairment. Worked somewhat less than expected and/or had mild difficulties in carrying out housework.
4 = Moderate impairment. Has missed a lot of housework when expected and/or has had considerable difficulties in carrying out housework.
5 = Severe impairment. Has missed a great deal of housework when expected to work and/or has been virtually unable to carry out housework when he attempts it.
6 = No information.

(1*c*) Student: _____
Which of the following categories best characterizes the degree to which the patient's current school work has been impaired as a result of psychopathology?
0 = Not applicable. Because not currently enrolled in a student program for reasons other than psychopathology.
1 = No impairment – high level. Worked as much as would be expected if not symptomatic and got high grades.

2 = No impairment – satisfactory level. Worked as much as would be expected if not symptomatic and got satisfactory grades.
3 = Mild impairment. Worked somewhat less and/or got grades somewhat below expected if not symptomatic.
4 = Moderate impairment. Missed a lot of school work and/or got grades consistently below expected.
5 = Severe impairment. Missed most of school work and/or dropped out of school or got grades far below those expected.
6 = No information.

(1) Work (maximum of 1*a*, 1*b* and 1*c*): _____

(2) Interpersonal relations
Which of the following best characterizes the patient's level of interpersonal relationships with his family currently (past month)? [Provide separate ratings for spouse (2*a*), children (2*b*) and other relatives (2*c*).]

(2*a*) Interpersonal relations with spouse: _____
(2*b*) Interpersonal relations with children: _____
(2*c*) Interpersonal relations with other relatives: _____
0 = Not applicable because does not have relatives in this category.
1 = Very good. Experiences very good relationships with this (these) family member(s), with only transient friction which is rapidly resolved. Feels only very minor or occasional need to improve quality of relationship, which is usually close and satisfying.
2 = Good. Argues occasionally, but arguments usually resolve satisfactorily within a short time. May occasionally prefer not to be with them because of dissatisfaction with them or be actively working with them to improve relationship.
3 = Fair. Often argues with this (these) family member(s) and takes a long time to resolve arguments. May withdraw from this person (these people) due to dissatisfaction. Often thinks that relationship needs to be either more harmonious or closer emotionally even when no conflict is present. For those relatives not living with the subject, contacts with them by choice are less frequent than feasible or rarely enjoyed very much when made.
4 = Poor. Regularly argues with this (these) family member(s) and such arguments are rarely ever resolved satisfactorily. Regularly prefers to avoid contact with them and/or feels great deficit in emotional closeness. For those family members out of the household, subject avoids seeing them as much as possible and derives no pleasure from contact when made.
5 = Very poor. Either constantly argues with this (these) family member(s) or withdraws from them most of the time. Separated or divorced from spouse or children moved out of household or almost always hostile to them when in contact.
6 = Variable. Different levels for various members of this group, and would warrant a rating of good or better (2, 1) with at least 1 member of this group. (Rate as 2.)
7 = Variable. Different levels for various members of this group, and would not warrant a rating of good or better (2, 1) with any member of this group. (Rate as 4.)
8 = No information.

(2*d*) Interpersonal relations with friends: _____
Which of the following best characterizes the patient's interpersonal relationships with friends currently (past month)?
1 = Very good. Had several special friends that he saw regularly and frequently and was close to.
2 = Good. Had at least two special friends that he saw from time to time and was fairly close to.
3 = Fair. Had only one special friend that he saw from time to time and was fairly close to; or contacts limited to several friends that he was not very close to emotionally.
4 = Poor. Had no special friends he saw from time to time and was fairly close to; or contacts limited to one or two friends that he was not very close to.
5 = Very poor. Had no special friends and practically no social contacts.
6 = No information.

(2) Interpersonal relations (maximum of 2*a*, 2*b*, 2*c* and 2*d*): _____

(3) Satisfaction: _____
Which of the following best characterizes the patient's overall level of satisfaction (contentment, degree to which he feels fulfilled, gratification derived from activities) for the past week.
1 = Very good. Transient problems may occur, but generally satisfied with all aspects of his life.
Occasional minor dissatisfaction in one area, but overall is quite content with himself, job, family, friends, activities and finances.
2 = Good. Mild dissatisfaction persists, but only in one area or is intermittent in several areas. In balance, is generally content and able to enjoy life most of the time, but does think there should be some improvement in either occupational role, interpersonal relations, sexual activities or finances.
3 = Fair. Moderate dissatisfaction in one or more areas, which is relatively persistent. Either discontent with occupational role, interpersonal relations, sexual activities or finances.
4 = Poor. Very dissatisfied in most areas and derives little pleasure from life. Rarely able to derive any satisfaction from activities or relationships.
5 = Very poor. Derives no satisfaction from anything. May feel no desire to carry out the smallest task or to be with other people.
6 = No information.

(4) Recreation: _____
At what level has the patient been involved in and able to enjoy recreational activities and hobbies (reading, spectator or participant sports, gardening, music, sewing, attending parties or gatherings, church or community organizations) in the past week.
1 = Very good. Has at least two activities which he enjoys fully and frequently.
2 = Good. Participates in several activities and does not always enjoy them fully; or participates in fewer activities or less frequently than optimal, but enjoys participation.

3 = Fair. Occasional participation in recreational activities or hobbies; or limited enjoyment when participation occurs.
4 = Poor. Some participation in recreational activities or hobbies, and derives very little enjoyment from such activities.
5 = Very poor. No involvement in recreational activities or hobbies.
6 = No information.

SUMMARY
(1) Work (maximum of 1*a*, 1*b* and 1*c*): _____
(2) Interpersonal relations (maximum of 2*a*, 2*b*, 2*c* and 2*d*): _____
(3) Satisfaction: _____
(4) Recreation: _____
Total score (sum of 1, 2, 3 and 4): _____

RIFT Scoring Template

Prepared by: Todd Dunn M.S., Abu Minhajuddin Ph.D., and Margaret Marcotte M.S.

Variable Names: LFRFTT = LIFE –RIFT Total Score
Imwrkt = Occupational Functioning: Item 1
Imprlt = Interpersonal Functioning: Item 2
GLBSAT = Global Satisfaction: Item 3
REC = Recreational Functioning: Item 4

OPSCAN:

Version date: None

Number of Items scored: 4

Scale: 4 (no functional impairment) – 20 (severe functional impairment)

Protocol: Prophylactic Cognitive Therapy for Recurrent Depression

Different Item Weights: No

The RIFT provides a psychometrically valid approach to scoring the LIFE Psychosocial Interview (LIFE-PI). Since the LIFE-PI assesses psychosocial functioning over a four-month span, the LIFE-PI can produce four RIFT total scores, one total score for each month assessed. The months being assessed are determined by the PHASE and WK variables. PHASE indicates which phase of the study being assessed, either acute or post randomization (Only acute phase data is included in this study.). The WK variable indicates the starting month of the data being collected. Only WK 1 is used in this study, which refers to the four months of the acute phase. For example, with PHASE = 1 (acute phase) and WK = 1 (month 1), then the months being assessed by the RIFT are month 1 (diagnostic evaluation or baseline assessment) and months 2, 3, and 4, which are the three months of A-CT.

The following items are scored as following:

PHASE
1= Acute Phase
2= Post-Randomization

WK
1 = week 1
2 = week 5
3 = week 9
4 = week 13

5 = week 17
6 = week 21
7 = week 25
8 = week 29

Impwrk_1 to Impwrk_4
0 = Not applicable (did not work for other reasons)
1 = No impairment, high level
2 = No impairment, satisfactory level
3 = Mild impairment
4 = Moderate impairment
5 = Severe impairment
6 = No information

Imphs_1 to Imphs_4
0 = Not applicable (did not work for other reasons)
1 = No impairment, high level
2 = No impairment, satisfactory level
3 = Mild impairment
4 = Moderate impairment
5 = Severe impairment
6 = No information

Imstwk_1 to Imstwk_4
0 = Not applicable (did not work for other reasons)
1 = No impairment, high level
2 = No impairment, satisfactory level
3 = Mild impairment
4 = Moderate impairment
5 = Severe impairment
6 = No information

Imwrkt_1 to Imwrkt_4
Enter the maximum observed score from Impwrk, Imphs and Imstwk (only use numbers 1 through 5)

Mate_1 to Mate_4
0= Not applicable (no spouse/mate)
1 = Very Good
2 = Good
3 = Fair
4 = Poor
5 = Very poor
6 = Variable, at least one mate rates 1 or 2 (rate as 2)
7 = Variable, no mate rates 1 or 2 (rate as 4)

8 = No information

Child_1 to Child_4
0 = Not applicable (no children)
1 = Very Good
2 = Good
3 = Fair
4 = Poor
5 = Very poor
6 = Variable, at least one child rates 1 or 2 (rate as 2)
7 = Variable, no child rates 1 or 2 (rate as 4)
8 = No information

Other_1 to Other_4
0 = Not applicable (no other important relatives)
1 = Very Good
2 = Good
3 = Fair
4 = Poor
5 = Very poor
6 = Variable, at least one relative rates 1 or 2 (rate as 2)
7 = Variable, no relative rates 1 or 2 (rate as 4)
8 = No information

Friend_1 to Friend_4
1 = Very Good
2 = Good
3 = Fair
4 = Poor
5 = Very poor
6 = No information

Imprlt_1 to Imprlt_4
Enter the maximum observed score from Mate, Child, Other and Friend (only use numbers 1 through 5)

GLBSAT_1 to GLBSAT_4
1 = Very Good
2 = Good
3 = Fair
4 = Poor
5 = Very poor

REC_1 to REC_4
1 = Very Good

2 = Good
3 = Fair
4 = Poor
5 = Very poor

LFRFTT_1 to LFRFTT_4
Sum the scores for Imwrkt, Imprlt, GLBSAT and REC

Original Reference:

Leon, A. C., Solomon, D. A., Mueller, T. I., Turvey, C. L., Endicott, J., & Keller, M. B. (1999). The Range of Impaired Functioning Tool (RIFT): A brief measure of functional impairment. *Psychological Medicine, 29,* 869-878.

Logic of Scoring Algorithm: To score Items 1 and 2, you take the highest score, up to and not exceeding 5, observed on each Item's respective queries. To score Items 3 and 4, you simply score the single query on each Item. To calculate the RIFT Total Score, you sum the scores from the four Items.

Missing Data: Items are not to be prorated. If missing data prevents the scoring of one item, you should not compute the total score. Scores of 0 and 6 are coded as missing and are not included in Item scores (p. 871; Leon et al., 1999).

APPENDIX G

Sub-analysis of Preliminary Aim 1: Investigating the Psychometric Quality of Measures Included on the Index of Psychosocial Functioning

Sub-analysis of Preliminary Aim 1: Investigating the Psychometric Quality of Measures Included on the Index of Psychosocial Functioning

Research Questions, Hypotheses, and Statistical Analyses

Research Question: What is the psychometric quality of the RIFT and SAS-SR compared to the DYS, GAF, and Q-LES-Q?

Rationale: In order to create an acceptable index of psychosocial functioning, the measures included on the index, which in this study are the RIFT and SAS-SR, much be psychometrically valid. Specifically, the RIFT and SAS-SR were required to demonstrate: a) internal consistency reliability, b) convergent and discriminant validity, and c) sensitivity to pre- to post-A-CT changes in psychosocial functioning.

Research Questions Operationalized:

1) Do total scores from the DYS, RIFT, SAS-SR, and Q-LES-Q have acceptable internal consistency reliability coefficients (i.e., $\alpha \geq .80$)?
2) Do total scores from the DYS, GAF, RIFT, and SAS-SR show convergent validity by having significant, positive correlations with each other at baseline?
3) Do the DYS, RIFT, SAS-SR, and Q-LES-Q show convergent validity by having significant, positive correlations between similar constructs across different measures at baseline (i.e., correlations coded "A" on Table G3)?
4) Referring to Table G3, do the DYS, RIFT, SAS-SR, and Q-LES-Q show discriminant validity by having, on average:

a. Higher "A" correlations than "C" correlations at baseline?

b. Higher "A" correlations than "B" correlations at baseline?

c. And higher "B" correlations than "C" correlations at baseline?

5) Do total scores from the DYS, GAF, RIFT, and SAS-SR show discriminant validity by not being highly correlated (i.e., $r < .50$; Cohen, 1988) with total scores from the BDI, HRSD, and IDS-SR?

6) Do total scores from the DYS, GAF, RIFT, SAS-SR, and Q-LES-Q show sensitivity to pre- to post-A-CT changes in psychosocial functioning, as measured by standardized mean gain effect sizes?

Hypotheses:

1) Based on previous research (e.g., Leon et al., 1999; Weissman et al., 1976), the RIFT and SAS-SR were expected to have acceptable levels of internal consistency reliability.

2 – 5) No hypotheses were generated, as these were exploratory analyses

6) Based on our quantitative review of the literature, it was predicted that the GAF and SAS-SR would have large effect sizes ($d > 0.80$), the RIFT and Q-LES-Q medium effect sizes ($0.50 < d < 0.80$), and the DYS a small-to-medium effect size ($0.20 < d < 0.50$).

Statistics: To test hypothesis one, Cronbach's Alphas were calculated for total scores from the DYS, RIFT, SAS-SR, and Q-LES-Q that corresponded with treatment baseline and the beginning, middle, and end of A-CT. Specifically, DYS and SAS-SR total scores were used from the diagnostic evaluation or follow-up, week one and seven of A-CT, and the post-A-CT blind evaluation (see Table 8, p. 80). Total scores from the

diagnostic follow-up and post-A-CT blind evaluation were used from the Q-LES-Q, and retrospective scores from the RIFT taken at the post-A-CT blind evaluation were used that represented the month of diagnostic evaluation and the first, second, and third month of A-CT (see Table 8). Burlingame et al. (1995) suggested that internal consistency reliability estimates be ≥ 0.80 for outcome instrumentation.

To test hypothesis two, Pearson product-moment correlation coefficients were used to determine the degree to which total scores from the DYS, GAF, RIFT, and SAS-SR covaried at treatment baseline. Baseline assessments occurred during the diagnostic evaluation or diagnostic follow-up for the DYS, GAF, and SAS-SR, and retrospectively during visit 101 for the RIFT (see Table 8).

To test hypotheses three and four, a multitrait-multimethod matrix was created using Pearson product-moment correlation coefficients (see Table G3). Only baseline total scores or subscales were used that measured domains of psychosocial functioning, including the total score from the DYS and pertinent subscales from the RIFT, SAS-SR, and Q-LES-Q. This was done because total scores representing "overall" psychosocial functioning do not measure a specific trait or domain of the construct. Besides, the agreement between total scores representing overall psychosocial functioning was addressed in hypothesis two (see above paragraph).

In order for measures of a particular construct to have convergent validity, correlations of similar traits across different methods need to be "significantly different from zero and sufficiently large to encourage further examination of validity" (p. 82; Campbell & Fiske, 1959). Similarly, in order to have discriminate validity, three conditions must exist: a) correlations of similar traits across different methods are higher

than correlations of different traits across different methods, b) correlations of similar traits across different methods are higher than correlations of different traits across similar methods, and c) correlations of different traits across similar methods is higher than correlations of different traits across different methods. To facilitate comparison of correlations across constructs and measures, the absolute value of pertinent correlations was averaged, producing overall correlation coefficients that were easily comparable. The absolute value of each correlation coefficient was used because the magnitude of the relationship was of interest rather than the direction. This process created three overall correlation coefficients that coincided with the three types of correlations discussed by Campbell et al.

To test hypothesis five, Pearson product-moment correlation coefficients were used to determine the degree to which total scores from the DYS, GAF, RIFT, and SAS-SR covaried with total scores from the BDI, HRSD, and IDS-SR. Only total scores from baseline assessments were used (see Table 8). Cohen (1988) categorized a correlation of 0.50 as large, 0.30 as moderate, and 0.10 as small.

To test hypothesis 6, pre- to post-A-CT total scores on the DYS, GAF, RIFT, SAS-SR, and Q-LES-Q were converted into standardized mean gains (ES_{mg}), such that:

$$ES_{mg} = \frac{M_{T1} - M_{T2}}{\sqrt{(s_{T1}^2 + s_{T2}^2) / 2}}$$

In this equation, m_{t1} and s_{t1}^2 and m_{t2} and s_{t2}^2 equaled the means and variances of pre- to post-A-CT total scores of psychosocial functioning, respectively. Using conventional norms (e.g., Cohen, 1988), the values of effect sizes were compared across instruments to

quantify the treatment effects of A-CT on psychosocial functioning, such that 0.80, 0.50, and 0.20 represented large, medium, and small effects, respectively.

Results

It was proposed that total scores from the RIFT and SAS-SR be standardized and averaged to form an index of psychosocial functioning, which would subsequently be entered into the structural equation model in the primary aim. This sub-analysis of preliminary aim one focused on evaluating the RIFT and SAS-SR's internal consistency reliability, convergent and discriminant validity, and sensitivity to pre- to post-A-CT changes. Below, results were presented for the RIFT and SAS-SR, as well as for the DYS, GAF, and Q-LES-Q for comparison. While not perfect, the RIFT and SAS-SR demonstrated adequate psychometric quality, justifying their inclusion in the index of psychosocial functioning.

Do total scores from the DYS, RIFT, SAS-SR, and Q-LES-Q have acceptable internal consistency reliability coefficients (i.e., $\alpha \geq .80$)?

As seen in Table G1, internal consistency reliability coefficients for the RIFT and SAS-SR did not reach the 0.80 level recommended for instruments used to evaluate treatment outcomes (Burlingame et al., 1995), while coefficients for the DYS and Q-LES-Q did. To assist in the interpretation of these results, three issues were explored.

Table G1

Internal Consistency Reliability of Measures of Psychosocial Functioning and Depressive Symptoms Severity during Diagnostic Evaluation and A-CT

		Internal Consistency Coefficients							
		Diagnostic Evaluation		First Session of A-CT		Psychoeducation Visit or Week 7 of A-CT		Post-A-CT Blind Evaluation	
Measure	Items	alpha	s^2	alpha	s^2	alpha	s^2	alpha	s^2
BDI	21	0.83	81.20	0.86	83.47	0.91	82.17	0.92	80.30
DYS	32	0.94	544.61	0.95	561.13	0.95	578.43	0.95	587.03
HRSD-17	17	0.51	17.10	0.63	23.36	0.84	40.47	0.83	41.00
IDS-SR	28	0.82	107.57	0.83	110.32	0.91	148.00	0.92	131.60
RIFT [a]	4	0.60	7.50	0.65	7.96	0.71	8.97	0.75	9.76
SAS-SR	56	0.77	0.20	0.77	0.20	0.78	0.22	0.76	0.18
QOLESQ	93	0.92	2330.33					0.90	3677.30

Note. A-CT = Acute-phase Cognitive Therapy; BDI = Beck Depression Inventory; DYS = Dyadic Adjustment Scale; GAF = Global Assessment of Functioning Scale; HRSD-17 = Hamilton Rating Scale for Depression-17; IDS-SR = Inventory for Depressive Symptomatology – Self-Report; Q-LES-Q = Quality of Life Enjoyment and Satisfaction Questionnaire; RIFT = Longitudinal Interval Follow-up Evaluation – Range of Impaired Functioning Tool; SAS-SR = Social Adjustment Scale-Self Report; alpha = internal consistency reliability coefficient; s^2 = variance.

[a] RIFT was administered by an evaluator at visit 101, and scores correspond to the month of diagnostic evaluation and the first, second, and third month of A-CT.

First, it might have been the case that Burlingame et al.'s coefficient of 0.80 for outcome instrumentation was too stringent. For instance, Schmitt (1996) claimed that while the field of psychology most often used a cutoff of 0.70 to represent adequate internal consistency reliability, the process of evaluating internal consistency reliability was more complex than a single criterion. Instead, acceptable internal consistency reliability might best be determined on a case by case basis and interpreted in the context of other psychometric properties. Schmitt went on to report that instruments with internal consistency reliability coefficients as low as 0.50 could still be considered acceptable when they possess reasonable convergent validity and other relevant psychometric qualities.

Second, as the number of items increases on an instrument, Cronbach's alpha increases. Cortina (1993) reports that Cronbach's alpha is underestimated on instruments with less than 19 items, as "the relationship between number of items and alpha is curvilinear (Komorita & Graham, 1965) and begins to level off before the number of items reaches 19" (p. 101). The subscales of the SAS-SR and total scores from the RIFT and HRSD-17 have less than 19 items (see Table G1).

Third, internal consistency reliability coefficients may be underestimated when inter-item variation is restricted, which occurs when there is little variation in a data set. For example, across the initial diagnostic evaluation, first and seventh week of A-CT, and post-A-CT blind evaluation, the variance of HRSD-17 total scores equaled 17.10, 23.36, 40.47, and 41.00, respectively (see Table G1). The HRSD-17 was the only measure to be restricted early on in the study, as only patients with HRSD-17 total scores ≥ 14 at both diagnostic interviews were eligible to enter A-CT. As seen in Table G1, when scores on

the HRSD-17 were no longer restricted, variability in the dataset and internal constancy reliability coefficients increased.

In sum, total scores from the RIFT and SAS-SR did not meet the criterion prescribed for outcome questionnaires (Burlingame et al., 1995). Evidence was reviewed that suggested internal consistency reliability was: a) not solely determined by a single criterion and b) influenced by the number of items on an instrument and variation in the data set. When these factors were considered, it was concluded that the RIFT and SAS-SR possessed adequate internal consistency reliability.

Do total scores from the DYS, GAF, RIFT, and SAS-SR show convergent validity by having significant, positive correlations with each other at baseline?

The RIFT and SAS-SR could show convergent validity by being significantly correlated with each other and additional measures of psychosocial functioning, including the DYS and GAF. As seen in Table G2, total scores from the SAS-SR, RIFT, and GAF were significantly correlated, with r ranging from 0.26 (p = 0.01) to 0.38 (p = 0.01). Also, while the DYS did not significantly correlate with the RIFT or GAF, it did significantly correlate with the SAS-SR (r = 0.39, p = 0.01). All significant correlations were moderate in size (Cohen, 1988).

Table G2

Correlation Coefficients between Total Score from the DYS, GAF, RIFT, and SAS-SR

Measures	Correlation Coefficient		
	DYS	GAF	RIFT
GAF	0.11		
RIFT	-0.14	-0.26*	
SAS-SR	-0.39*	-0.38*	0.33*

Note. DYS = Dyadic Adjustment Scale; GAF = Global Assessment of Functioning Scale; RIFT = Longitudinal Interval Follow-up Evaluation – Range of Impaired Functioning Tool; SAS-SR = Social Adjustment Scale-Self Report.

* $p = 0.01$

Taken together, the GAF, RIFT, and SAS-SR showed moderate levels of convergent validity, suggesting these measures agreed somewhat in their assessment of overall psychosocial functioning but not so much as to indicate they were redundant. The DYS, perhaps due to its focus on dyadic functioning, only showed moderate levels of convergent validity with the SAS-SR, which includes four subscales assessing interpersonal functioning with extended family members, a spouse, children, and the family unit.

Do the DYS, RIFT, SAS-SR, and Q-LES-Q show convergent validity by having significant, positive correlations between similar traits across different methods of measurement at baseline?

Campbell et al. (1959) reported that instruments could show convergent validity by having significant correlations between subscales that measure similar traits across different methods of measurement. As seen in Table G3, all correlations between similar traits across different methods were significant at $p = 0.001$, averaging 0.32. Correlations ranged from 0.65, as observed between the DYS and Marital Subscale from the SAS-SR, to 0.24, as observed between the work item from the RIFT and Work Composite from the SAS-SR and the interpersonal item from the RIFT and DYS. Inter-subscale correlations were moderate to large in size (Cohen, 1988). According to these results, the RIFT and SAS-SR, as well as the DYS and Q-LES-Q, showed convergent validity when measuring domains specific to psychosocial functioning.

Table G3

Correlations between Total and Subscale Scores from Measures of Psychosocial Functioning at Baseline

		RIFT				SAS-SR				Q-LES-Q		
	Traits	IRL	WL	RecL	SatL	IRS	WS	RecS	IRQ	WQ	RecQ	SatQ
RIFT	IRL											
	WL	0.22**										
	RecL	0.29**	0.32**									
	SatL	0.19**	0.25**	0.41**								
SAS-SR	IRS	0.29**	0.05	0.03	0.03							
	WS	0.08	0.24**	0.09	0.14**	0.19**						
	RecS	0.22**	0.17**	0.27**	0.19**	0.25**	.23**					
Q-LES-Q	IRQ	-0.25**	-0.07	-0.21**	-0.14*	-0.26**	-0.32**	-0.47**				
	WQ	-0.10	-0.39**	-0.16**	-0.25**	-0.20**	-0.50**	-0.30**	0.40**			
	RecQ	-0.11*	-0.17**	-0.25**	-0.15**	-0.22**	-0.17**	-0.40**	0.44**	0.37**		
	SatQ	-0.17**	-0.25**	-0.27**	-0.29**	-0.36**	-0.42**	-0.44**	0.58**	0.65**	0.50**	
DYS		-0.24**	-0.07	-0.02	-0.03	-0.65**	-0.11	-0.21**	0.18**	0.10	0.24**	0.24**

Note. RIFT = Longitudinal Interval Follow-up Evaluation – Range of Impaired Functioning Tool; TSL = Total score from the RIFT; IRL = Interpersonal Relations Item from RIFT; WL = Work Item from RIFT; RecL = Recreational Item from RIFT; SatL = Satisfaction Item from RIFT; SAS-SR = Social Adjustment Scale-Self Report; TSS = Total score from the SAS-SR; IRS = Interpersonal Relationship Composite from SAS-SR; WS = Work Subscale from SAS-SR; RecS = Social/Leisure Subscale from SAS-SR; Q-LES-Q = Quality of Life Enjoyment and Satisfaction Questionnaire; IRQ = Social Relations Subscale from Q-LES-Q; WQ = Work Composite from the Q-LES-Q; RecQ = Leisure Subscale from the Q-LES-Q; SatQ = Satisfaction with General Activities Subscale from the Q-LES-Q; DYS = Dyadic Adjustment Scale; A = Correlation coefficients between similar constructs across different measures; B = Correlation coefficients between different constructs across the same measures; C = Correlation coefficients between different constructs across different measures.

* $p = 0.05$; ** $p = 0.01$

Do the DYS, RIFT, SAS-SR, and Q-LES-Q show Discriminant Validity by Meeting the Three Requirements of Discriminant Validity Outlined by Campbell et al. (1959)?

Referring to Table G4, Campbell et al. (1959) reported that instruments measuring a similar construct could show discriminant validity by having, on average, a) higher "A" correlations than "C" correlations, b) higher "A" correlations than "B" correlations, and c) higher "B" correlations than "C" correlations. After averaging the absolute value of correlations in each category (see Table G3), the DYS, RIFT, SAS-SR, and pertinent subscales from Q-LES-Q did not meet all three requirements of discriminant validity. When subscales assessing psychosocial functioning from the Q-LES-Q were removed from this analysis due to high intra-subscale correlations, however, the DYS, RIFT, and SAS-SR did show discriminant validity by meeting Campbell et al.'s three requirements.

Table G4

Description of the Correlations Used by Campbell et al. (1959) to Determine Discriminant Validity

Correlation Type	Trait and Method of Subscales being Correlated
A	Similar traits across different methods of measurement
B	Different traits across similar methods of measurement
C	Different traits across different methods of measurement

For the DYS, RIFT, SAS-SR, and Q-LES-Q, the averaged correlation coefficients between similar traits across difference methods ($r = 0.32$) and different traits across similar methods ($r = 0.35$) exceeded the averaged correlation coefficient between different traits across different methods ($r = 0.18$), satisfying two requirements of discriminant validity. However, the averaged correlation coefficient between similar traits across difference methods ($r = 0.32$) did not exceed the averaged correlation coefficient between different traits across similar methods ($r = 0.35$). According to these results, the domains of psychosocial functioning measured by the DYS, RIFT, SAS-SR, and Q-LES-Q appeared distinct across, but not within, instruments.

Upon further examination of Table G3, subscales from the Q-LES-Q were excluded from this analysis because their intra-subscale correlation coefficients were large (Cohen, 1988), averaging 0.49. After excluding the Q-LES-Q, the averaged correlation coefficients between similar traits across difference methods ($r = 0.34$) and different traits across similar methods ($r = 0.26$) exceeded the averaged correlation coefficient between different traits across different methods ($r = 0.10$), and the averaged correlation coefficient between similar traits across difference methods ($r = 0.34$) exceeded the averaged correlation coefficient between different traits across similar methods ($r = 0.26$). As a result, compared to the DYS and each other, the RIFT and SAS-SR showed acceptable levels of discriminant validity, as the domains of psychosocial functioning were sufficiently distinct both across and within instruments.

Do measures of psychosocial functioning (i.e., DYS, GAF, RIFT, and SAS-SR) show discriminant validity by not being highly correlated (i.e., ≥ 0.50) with measures of depressive symptoms severity (i.e., BDI, HRSD, and IDS-SR)?

The DYS, GAF, RIFT, and SAS-SR could also show discriminant validity by not highly correlating with measures of a different construct, like depressive symptom severity. As seen in Table G5, the DYS, GAF, and RIFT showed acceptable levels of discriminant validity by not highly correlating with the BDI, IDS-SR, and HRSD-17. The SAS-SR also showed acceptable levels of discriminant validity with the HRSD-17, but not with the BDI or IDS-SR. Specifically, while the SAS-SR highly correlated with the BDI (r = 0.64, p = 0.01) and IDS-SR (r = 0.64, p = 0.01), correlations of the DYS, GAF, and RIFT with measures of depressive symptom severity were small to moderate in size, ranging from -0.45 to -0.06 (see Table G5).

Table G5

Correlation between Measures of Depressive Symptom Severity (i.e., BDI, IDS-SR, and HRSD-17) and Psychosocial Functioning (i.e., DYS, GAF, RIFT, and SAS-SR)

Measures of Depressive Symptom Severity	Measures of Psychosocial Functioning			
	DYS	GAF	RIFT	SAS-SR
BDI	-0.23*	-0.36*	0.29	0.64*
HRSD-17	-0.06	-0.45*	0.22*	0.40*
IDS-SR	-0.20*	-0.35*	0.28*	0.64*

Note. BDI = Beck Depression Inventory; DYS = Dyadic Adjustment; GAF = Global Assessment of Functioning Scale; HRSD = Hamilton Rating Scale for Depression-17; IDS-SR = Inventory for Depressive Symptomatology – Self-Report; RIFT = Longitudinal Interval Follow-up Evaluation – Range of Impaired Functioning Tool; SAS-SR = Social Adjustment Scale-Self Report.

* $p = 0.01$

Overall, the RIFT showed acceptable discriminant validity with measures of depressive symptom severity. On the other hand, the SAS-SR only showed discriminant validity when compared to clinician-rated, but not self-report, measures of depressive symptom severity.

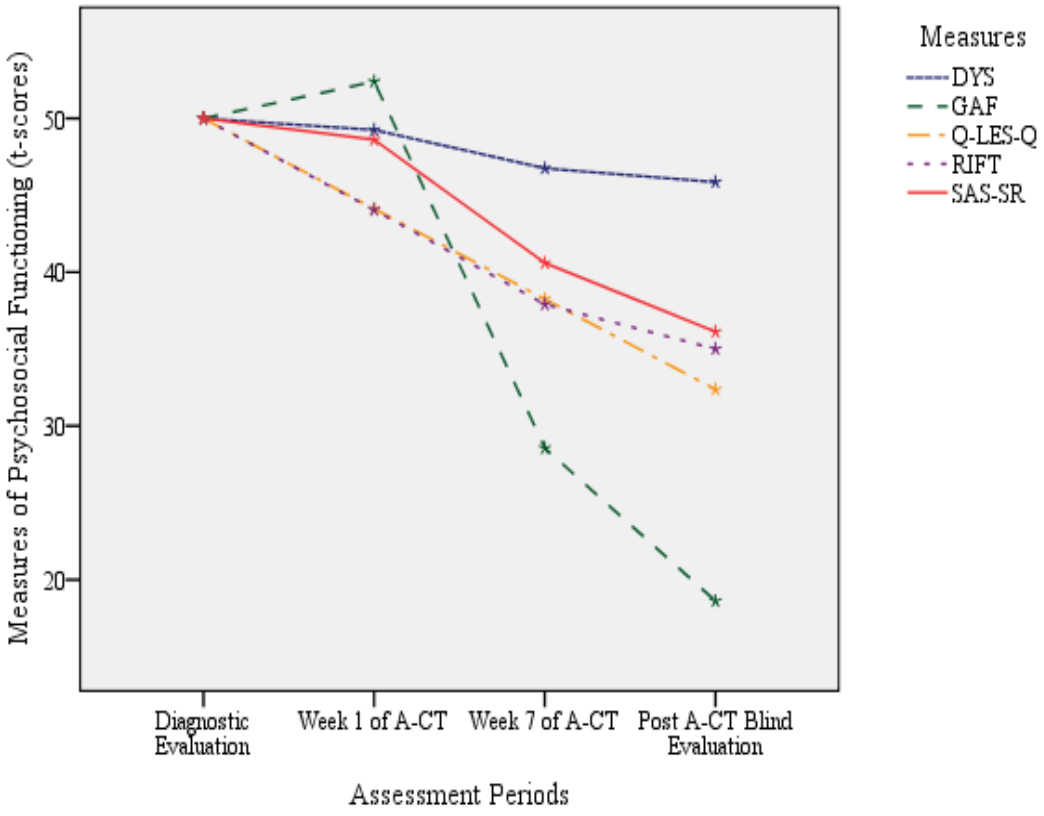
Do total scores from the DYS, GAF, RIFT, SAS-SR, and Q-LES-Q show sensitivity to pre- to post- A-CT changes in psychosocial functioning, as measured by their standardized mean gain effect sizes?

For this analysis, pre- to post-A-CT changes in DYS, GAF, RIFT, SAS-SR, and Q-LES-Q total scores were calculated using the standardized mean gain effect size. To calculate this effect size, pre- and post-A-CT total score means were subtracted and divided by their pooled variance. Based on a quantitative review of the literature (Dunn et al., 2007), it was predicted that the GAF and SAS-SR would have large effect sizes ($d > 0.80$), the RIFT and Q-LES-Q medium effect sizes ($0.50 < d < 0.80$), and the DYS a small-to-medium effect size ($0.20 < d < 0.50$). Results partially confirmed this hypothesis.

As predicted, large standardized mean gain effect sizes of 1.93 and 1.42 represented pre- to post-A-CT changes on the GAF and SAS-SR, respectively, and a small to medium standardized mean gain effect size ($d = 0.41$) represented the pre- to post-A-CT change on the DYS. Contrary to what was predicted, pre- to post-A-CT changes on the RIFT and Q-LES-Q equaled large standardized mean gain effect sizes of 1.41 and 1.57, respectively. These patterns of change were graphically represented in Figure G1, using T-scores. As seen in Figure G1, the RIFT and SAS-SR compared well with other commonly used measures of psychosocial functioning (e.g., DYS, GAF, Q-LES-Q), regarding sensitivity to changes in functioning that occurred during A-CT.

Figure G1

Changes in Psychosocial Functioning during Acute-Phase Cognitive Therapy for Five Measures of Psychosocial Functioning



Note. A-CT = Acute-phase Cognitive Therapy; DYS = Dyadic Adjustment Scale; GAF = Global Assessment of Functioning Scale; RIFT = Longitudinal Interval Follow-up Evaluation – Range of Impaired Functioning Tool; Q-LES-Q = Quality of Life Enjoyment and Satisfaction Questionnaire; SAS-SR = Social Adjustment Scale – Self-Report.

Discussion of Psychometric Quality of the RIFT and SAS-SR Compared to Other Measures of Psychosocial Functioning

This sub-analysis of preliminary aim one focuses on evaluating the psychometric property of measures included in the index of psychosocial functioning. While five measures of psychosocial functioning are employed in this study (i.e., DYS, GAF, RIFT, SAS-SR, and Q-LES-Q), the RIFT and SAS-SR represent the best fit for the index a-priori because they only cover material relevant to the construct and permit data collection for single individuals. The GAF and Q-LES-Q cover content outside the construct of psychosocial functioning (e.g., symptom severity, physical health, etc.), and the DYS only provides data on individuals in committed relationships. While research supports the psychometric quality of the RIFT (Leon et al., 1999) and SAS-SR (Weissman et al., 1976; Weissman, et al., 1978) individually, comparative studies investigating the psychometric quality of these and related instruments are rare (Vittengl et al., 2008; Weissman et al., 2001). Consequently, this sub-analysis: a) evaluates the RIFT and SAS-SR's internal consistency reliability, convergent and discriminant validity, and sensitivity to changes in psychosocial functioning from pre- to post-A-CT and b) compares these results to those from the DSY, GAF, and Q-LES-Q. Results not only serve to substantiate inclusion of RIFT and SAS-SR data on the index of psychosocial functioning, but they also fill a gap in the literature, informing the field on the comparative utility of the DYS, GAF, RIFT, SAS-SR, and Q-LES-Q in measuring psychosocial functioning in individuals receiving A-CT for MDD.

Results raise some question regarding the internal consistency reliability of the RIFT (median α = 0.68, range = 0.60 to 0.75) and SAS-SR (median α = 0.77, range = 0.76 to 0.78), as Cronbach's alpha coefficients for these instruments did not meet the 0.80 criterion recommended by Burlingame et al. (1995) for outcome instrumentation. However, a closer look at factors influencing Cronbach's alpha suggests the 0.80 criterion may be too stringent. In reality, an alpha coefficient as low as 0.50 can be acceptable when consideration is given to the degree of variation in the data set, number of items on the instrument, and other desirable psychometric qualities (Cortina, 1993; Schmitt, 1996). Therefore, instead of using a single criterion, acceptable internal consistency reliability might best be determined on a case by case basis and interpreted in the context of other psychometric properties.

Results are promising, when considering the convergent validity of the RIFT and SAS-SR. A multi-trait multi-method matrix that correlates items and subscales from the RIFT, SAS-SR, and Q-LES-Q, along with total scores from the DYS, shows indices of the same domain of psychosocial functioning significantly correlate, with an average correlation of 0.32. Only using total scores, the RIFT and SAS-SR also show acceptable levels of convergent validity, by significantly correlating with each other and the GAF and DYS (only the SAS-SR correlates significantly with the DYS). These correlations of convergent validity are moderate in size, suggesting there is significant agreement between these measures but not enough to indicate measures are redundant. Weissman et al (2001) and Vittengl et al. (2008) corroborate these results, showing the RIFT and SAS-SR demonstrate convergent validity with each other and the DYS, GAF, SASS, SF-36 Mental Component, and Inventory of Interpersonal Problems (Horowitz et al., 1988).

Findings related to discriminant validity are mixed for the RIFT and SAS-SR. The same multi-trait multi-method matrix shows the DYS, RIFT, and SAS-SR meet the three requirements of discriminant validity proposed by Campbell et al. (1959). Basically, these requirements ask that the DYS, RIFT, and SAS-SR have correlations between conceptually similar subscales that are higher than correlations between conceptually different subscales. If this is the case, then the measures are said to assess the different domains of psychosocial functioning with sufficient distinction, both across and within instruments.

The RIFT shows additional evidence of discriminant validity by not highly correlating (i.e., $\geq$ 0.50) with clinician-rated or self-report measures of depressive symptom severity (e.g., BDI, HRSD-17, and IDS-SR). However, while the SAS-SR does not highly correlate with the clinician-rated HRSD-17, it does highly correlate with self-report measures of depressive symptom severity, with *r* equaling 0.64 with the BDI and IDS-SR. This high correlation has been reported previously in the literature (e.g., Mulder et al., 2003) and may be attributed to the SAS-SR sharing significant content and/or method variance with the BDI and IDS-SR. For instance, Ro et al. (2007) suggest items on the SAS-SR are confounded because they use terms expressing negative affect when querying domains of psychosocial functioning (e.g., "Have you been *ashamed* [italics added] of how you do your work in the past two weeks?"). Consequently, individual's self-reported level of psychosocial functioning on the SAS-SR may also inadvertently reflect levels of their depressive symptomatology (e.g., depressed mood, excessive guilt or worthlessness, etc.).

Lastly, the RIFT and SAS-SR show sensitivity to changes in psychosocial functioning during A-CT. Specifically, large standardized mean gain effect sizes of 1.41 to 1.42 represent pre- to post-A-CT changes on the RIFT and SAS-SR, respectively. These effect sizes are almost identical to the effect size (d = 1.39) found in a meta-analysis of 14 studies investigating the degree to which psychosocial functioning improves from pre- to post-A-CT in individuals diagnosed with MDD (Dunn et al., 2007). The 14 studies included in this meta-analysis measure psychosocial functioning with the DYS, GAS, Q-LES-Q, SAS-SR, and a homemade scale of parental functioning.

Taken together, results show the RIFT and SAS-SR, while not perfect, assess the construct of psychosocial functioning with adequate reliability and validity. Also, the RIFT and SAS-SR appear to be sensitive enough to detect changes in psychosocial functioning during A-CT for MDD. Initial concerns over the internal consistency reliability of the RIFT and SAS-SR are alleviated by evidence suggesting: a) these instruments have convergent validity with each other and other measures of psychosocial functioning (Schmitt, 1996) and b) internal consistency reliability coefficients for the RIFT may be artificially low due to low number of items (i.e., four) and low variation in the dataset at baseline and the first month of A-CT (see Table G1). Additionally, concerns about the SAS-SR's discriminant validity due to its high correlation with the BDI and IDS-SR are somewhat alleviated because: a) the SAS-SR represents a "gold standard" in terms of being one of the most often used measures of psychosocial functioning in the literature on treatment for depression (e.g., Dunn et al., 2007; Hirschfeld et al., 2000); b) the SAS-SR has repeatedly measured psychosocial functioning in clinical trials for depression with acceptable levels of reliability and

validity (e.g., Hirschfeld et al., 2002; Imber et al., 1990; Mintz et al., 1992; Vittengl et al., 2004); and c) steps are taken in the primary aim to control for the covariation between measures of depressive symptom severity and psychosocial functioning. So, the possible limitations of the RIFT and SAS-SR in this preliminary investigation do not out weight their positive psychometric attributes or raise enough concern to warrant their removal from the index of psychosocial functioning.

APPENDIX H

Inter-Rater Reliability for the SCID-I Current MDE and HRSD-17 Across Study Sites

Project coordinators completed inter-site reliability studies on the HRSD-17 (see Table H1) and SCID-I Current MDE (see Table H2) to maintain high diagnostic reliability across sites. All clinicians and evaluators met quarterly to co-rate videotapes containing the HRSD-17 and SCID-I Current MDE interviews. Videotaped interviews were exchanged between sites, so two tapes from each site were rated each year.

Table H1

Inter-Rater Reliability between UT Southwestern and WPIC in HRSD-17 Ratings

HRSD-17	Both sites Mean (*SD*)	Dallas Mean (*SD*)	Pittsburgh Mean (*SD*)
Patient 1	11.6 (2.6)	11.3 (1.9)	12.0 (3.5)
Patient 2	13.3 (2.1)	14.0 (1.7)	12.8 (2.2)
Patient 3	16.3 (1.3)	17.2 (0.4)	15.8 (1.3)
Patient 4	18.0 (1.2)	18.6 (0.9)	17.6 (1.3)
Patient 5	2.1 (1.1)	1.5 (1.0)	2.3 (1.0)
Patient 6	23.3 (2.4)	23.6 (2.9)	23.1 (2.1)
Patient 7	10.5 (1.6)	10.6 (1.7)	10.5 (1.6)
Patient 8	16.8 (1.6)	16.0 (1.7)	17.6 (0.9)
Patient 9	16.1 (1.5)	16.0 (1.9)	16.2 (1.2)
Patient 10	15.0 (2.3)	15.2 (3.1)	14.8 (1.0)
Patient 11	15.2 (1.7)	15.9 (1.5)	13.8 (1.5)
Patient 12	14.5 (1.8)	15.0 (2.2)	13.8 (0.8)
Patient 13	10.6 (1.7)	11.0 (2.0)	10.4 (1.6)
Patient 14	13.8 (3.0)	11.2 (1.7)	15.1 (2.6)
Patient 15	15.4 (1.2)	15.4 (1.1)	15.3 (1.4)
Patient 16	17.4 (1.8)	17.7 (2.0)	17.1 (1.6)
Patient 17	28.1 (2.4)	26.6 (2.1)	29.2 (2.0)
Patient 18	11.3 (1.3)	11.2 (1.2)	11.4 (1.5)
Patient 19	16.1 (2.0)	15.3 (2.6)	16.6 (1.5)
Patient 20	11.8 (1.2)	12.0 (1.2)	11.7 (1.3)
Patient 21	21.9 (1.2)	22.0 (2.0)	21.8 (0.8)
Patient 22	7.6 (2.9)	7.9 (2.5)	7.3 (3.6)
Patient 23	12.3 (1.7)	12.1 (1.7)	12.5 (1.7)
Intraclass Correlation	0.72 [a] /0.99 [b]	0.63 [a] /0.98 [b]	0.80 [a] /0.99 [b]

[a] The Intraclass Correlation for individual raters across the videos. Both the limited number of raters and the heterogeneous mix of raters across the videos produce a very conservative reliability. [b] The Intraclass Correlation across the average of the raters for the videos. This approach limits the impact of the rater heterogeneity and shows the reliability if more raters of equivalent skill had been included.

Table H2

Inter-Rater Reliability of SCID-I Current Major Depressive Episode Ratings at UT Southwestern and WPIC

	Both sites	UT Southwestern	WPIC
SCID MDE Current	MDD? Yes/No	MDD? Yes/No	MDD? Yes/No
Patient 1	0/15	0/5	0/10
Patient 2	0/18	0/6	0/12
Patient 3	0/13	0/7	0/6
Patient 4	3/17	1/7	2/10
Patient 5	9/8	3/6	6/2
Patient 6	15/1	8/1	7/0
Patient 7	0/17	0/9	0/8
Patient 8	8/6	4/3	4/3
Patient 9	5/11	2/7	3/4
Patient 10	16/0	6/0	10/0
Patient 11	0/14	0/5	0/9
Patient 12	0/21	0/7	0/14
Patient 13	4/11	2/4	2/7
Patient 14	0/19	0/6	0/13
Patient 15	0/12	0/6	0/6
Patient 16	5/12	2/6	3/6
Average percent agreement:	87.6	86.6	88.6

APPENDIX I

Cognitive Therapy Scale Ratings

To assure therapist competence and adherence, each of the 15 therapists in this study completed extensive training in cognitive therapy and achieved and maintained Cognitive Therapy Scale (CTS; Young et al., 1980) scores ≥ 40 in order to treat study patients. From a total of 139 sessions that were rated, only 19 (14%) sessions, spread across eight therapists, fell below 40, and competency scores averaged 45.4 (*SD* = 5.9) (see Table I1). Also, CTS scores for therapists from UT Southwestern (see Tables I2) and WPIC (see Tables I3) averaged > 40.

Table I1

Annual Cognitive Therapy Scale Scores across Sites (UT Southwestern and WPIC) and Therapists

	Year								Total
Measures	1	2	3	4	5	6	7	8	Years 1-8
Mean	49.7	44.3	44.2	46.4	44.8	45.9	46.7	45.5	45.4
SD	9.3	7.1	5.5	5.0	5.2	4.6	4.5	6.5	5.9
# of Sessions Rated	9	15	25	10	27	30	12	11	139
Sessions Rated < 40	0	4	5	1	5	2	1	1	19

Table I2

Cognitive Therapy Scale Scores for Therapists from UT Southwestern

Measures	Therapist ID#	Year 1	2	3	4	5	6	7	8	Total Years 1-8
Mean	1	46.5	37.8	41.1						40.4
SD		4.5	3.2	5.3						5.1
# of Sessions Rated		1	2	3						6
Sessions Rated < 40		0	2	1						3
Mean	2	40.8	46.0	41.1	43.6	47.6	45.6	45.1	43.2	44.1
SD		2.3	3.1	1.9	2.3	5.0	2.4	2.1	7.5	4.5
# of Sessions Rated		1	1	5	2	5	4	2	3	23
Sessions Rated < 40		0	0	2	0	0	0	0	1	3
Mean	3	47.5	43.1							44.7
SD		3.3	2.5							3.5
# of Sessions Rated		1	2							3
Sessions Rated < 40		0	1							1
Mean	4		42.4	44.7						43.0
SD			7.1	1.2						6.1
# of Sessions Rated			2	1						3
Sessions Rated < 40			1	0						1
Mean	5		45.3	41.3	45.3	42.3				43.7
SD			2.9	1.5	3.8	3.2				3.4
# of Sessions Rated			2	1	3	4				10
Sessions Rated < 40			0	0	0	1				1
Mean	6		45.9	42.2	49.5	45.3	45.9	45.1	47.3	45.4
SD			3.9	3.3	7.3	3.4	3.3	2.8	1.8	3.8
# of Sessions Rated			2	3	2	4	5	5	2	23
Sessions Rated < 40			0	2	0	0	0	0	0	2
Mean	7						45.1	46.3	45.1	45.2
SD							1.9	0.6	3.8	2.5
# of Sessions Rated							7	1	3	11
Sessions Rated < 40							0	0	0	0
Mean	Annual Site Totals	44.6	42.8	41.6	45.8	44.9	45.5	45.3	44.9	45.5
SD		4.4	5.0	3.3	4.8	4.4	2.6	2.4	5.5	5.8
# of Sessions Rated		3	11	13	7	13	16	8	8	79
Sessions Rated < 40		0	4	5	0	1	0	0	1	11

Table I3

Cognitive Therapy Scale Scores for Therapists from WPIC

Measures	Therapist ID#	Year 1	2	3	4	5	6	7	8	Total Years 1-8
Mean				50.5	42.3	38.8	46.8			45.0
SD	8			2.5	4.5	3.7	7.9			7.0
# of Sessions Rated				3	1	4	5			13
Sessions Rated < 40				0	1	4	1			6
Mean				56.0		45.7	46.7			48.1
SD	9			2.8		4.7	8.3			7.4
# of Sessions Rated				2		3	5			10
Sessions Rated < 40				0		0	1			1
Mean			63.3	52.5	52.0			51.5	52.6	57.5
SD	10		0.6	2.1	1.7			4.8	7.2	5.3
# of Sessions Rated			3	1	1			2	2	9
Sessions Rated < 40			0	0	0			0	0	0
Mean		59.5	61.0	53.0		54	59	38		55.4
SD	11	8.4	0.0	2.6		0	0	0		8.1
# of Sessions Rated		4	1	3		1	1	1		11
Sessions Rated < 40		0	0	0		0	0	1		1
Mean		63.5			49.0	57.0	47.0			54.8
SD	12	0.7			5.7	0	0			8.0
# of Sessions Rated		2			1	1	1			5
Sessions Rated < 40		0			0	0	0			0
Mean				50.5		47.5				49.0
SD	13			0.7		10.6				6.4
# of Sessions Rated				2		2				4
Sessions Rated < 40				0		0				0
Mean				47.0						47.0
SD	14			0.0						0.0
# of Sessions Rated				1						1
Sessions Rated < 40				0						0
Mean						44.7	42	44	44	43.8
SD	15					4.7	2.8	0	4.2	3.2
# of Sessions Rated						3	2	1	1	7
Sessions Rated < 40						0	0	0	0	0
Mean	Annual Site Totals	60.8	62.8	51.9	47.6	44.7	46.9	49.0	53.2	49.9
SD		6.8	1.3	3.0	5.7	7.0	7.6	9.9	9.5	8.3
# of Sessions Rated		6	4	12	3	14	14	4	3	60
Sessions Rated < 40		0	0	0	1	4	2	1	0	8

BIBLIOGRAPHY
(* studies included in the meta-analysis)

Aberg-Wistedt, A., Agren, H., Ekselius, L., Bengtsson, F., & Akerblad, A. C. (2000). Sertraline versus paroxetine in major depression: Clinical outcome after six months of continuous therapy. *Journal of Clinical Psychopharmacology, 20,* 645-652.

Abramson, L. Y., Seligman, M. E. P., & Teasdale, J. (1978). Learned helplessness in humans: Critique and reformulation. *Journal of Abnormal Psychology, 87,* 49-74.

Agosti, V., & Stewart, J. W. (1998). Social functioning and residual symptomatology among outpatients who responded to treatment and recovered from major depression. *Journal of Affective Disorders, 47,* 207-210.

Ajzen, I. (1991). The theory of planned behavior. *Organizational Behavior and Human Decision Processes, 50,* 179-211.

American Psychiatric Association. (1980). *Diagnostic and Statistical Manual of Mental Disorders* (3rd ed.). Washington, DC: American Psychiatric Association.

American Psychiatric Association. (1994). *Diagnostic and Statistical Manual of Mental Disorders* (4th ed.). Washington, DC: American Psychiatric Association.

American Psychiatric Association. (2000). *Diagnostic and Statistical Manual of Mental Disorders DSM-IV-TR* (Text Revision). Washington, DC: American Psychiatric Association.

Ansseau, M., Dierick, M., Buntinkx, F., Cnockaert, P., De Smedt, J., Van Den Haute, M., et al., (2004). High prevalence of mental disorders in primary care. *Journal of Affective Disorders, 78,* 49-55.

Arean, P. A., Ayalon, L., Hunkeler, E., Lin, E. H., Tang, L., Harpole, L., Hendrie, H., Williams, J. W., Unutzer, J., & IMPACT Investigators. (2005). Improving depression care for older, minority patients in primary care. *Medical care, 43,* 381-390.

Arnow, B. A., Blasey, C., Manber, R., Constantino, M. J., Markowitz, J. C., Klein, D. N., Thase, M. E., Kocsis, J. H., & Rush, A. J. (2007). Dropouts versus completers among chronically depressed outpatients. *Journal of Affective Disorders, 97,* 197-202.

Bagby, R. M., Quilty, L. C., Segal, Z. V., McBride, C. C., Kennedy, S. H., & Costa, P. T. (2008). Personality and differential treatment response in major depression: A randomized controlled trial comparing cognitive-behavioural therapy and pharmacotherapy. *Canadian Journal of Psychiatry, 53*, 361–370.

Baker, R., & Hall, J. N. (1988). REHAB: a new assessment instrument for chronic psychiatric patients. *Schizophrenia Bulletin, 14,* 95-113.

Bakish, D. (2001). New standard of depression treatment: Remission and full recovery. *Journal of Clinical Psychiatry, 62 (Supplement 26)*, 5-9.

Bandura, A. (1962). Social learning through imitation. In M. R. Jones (Ed.), *Nebraska Symposium on Motivation* (pp. 211-269). Lincoln, Nebraska: University of Nebraska Press.

Bandura, A. (1986). *Social Foundations of Thought and Action: A Social Cognitive Theory.* Englewood Cliffs, NJ: Prentice Hall.

Bandura, A. (1997). *Self-Efficacy: The Exercise of Control.* New York: W. H. Freeman and Company

Bandura, A. (1999). Social cognitive theory of personality. In L. A. Pervin & O. P. John (Eds.), *Handbook of Personality: Theory and Research* (pp. 154-196). New York: Guilford Press.

Bandura, A. (2001). Social cognitive theory: An agentic perspective. *Annual Review of Psychology, 52,* 1-26.

Baranowski, T., Perry, C. L., & Parcel, G. S. (1997). How individuals, environments, and health behavior interact: Social cognitive theory. In K. Glanz, F. M. Lewis, & B. K. Rimer (Eds.), *Health Behavior and Health Education: Theory, Research, and Practice* 2nd ed. (pp. 153-178). San Francisco: Jossey-Bass Publishers.

Baron, R. M., & Kenny, D. A. (1986). The Moderator-Mediator Variable Distinction in Social Psychological Research: Conceptual, Strategic, and Statistical Considerations. *Journal of Personality and Social Psychology, 51,* 1173-1182.

Barry, E. S., Naus, M. J., & Rehm, L. P. (2004). Depression and implicit memory: Understanding mood congruent memory bias. *Cognitive Therapy and Research, 28,* 387-414.

*Beach, S. R. H., & O'Leary, K. D. (1986). The treatment of depression occurring in the context of marital discord. *Behavior Therapy, 17*, 43-49.

Beach, S. R. H., & O'Leary, K. D. (1992). Treating depression in the context of marital discord: Outcome and predictors of response of marital therapy versus cognitive therapy, *Behavior Therapy,* 23, 507-528.

Bech, P. (2005). Social functioning: Should it become an endpoint in trials of antidepressants? *CNS Drugs, 19,* 313-324.

Beck, A. T., Rush, A. J., Shaw, B. F., & Emery, G. (1979). *Cognitive Therapy of Depression.* New York: Guilford.

Beck, A. T., Steer, R. A, & Brown, G. K. (1996). *Manual for the Beck Depression Inventory-II.* San Antonio, TX: Psychological Corporation.

Beck, A. T., Steer, R. A., & Garbin, M. G. (1988). Psychometric properties of the Beck Depression Inventory: Twenty-five years of evaluation. *Clinical Psychology Review, 8,* 77-100.

Beck, A.T., Ward, C. H., Mendelson, M., Mock, J. E., & Erbaugh, J. K. (1961). An inventory for measuring depression. *Archives of General Psychiatry, 4,* 561-571.

Beck, A. T., & Weishaar, M. (1993). Cognitive therapy. In A. Freeman, K. M. Simon, L. E. Beutler, & H. Arkowitz (Eds.), *Comprehensive Handbook of Cognitive Therapy* (pp. 21-36). New York: Plenum Press.

Bemporad, J. (1985). Long-term analytic treatment of depression. In E. E. Beckham & W. R. Leber (Eds.), *Handbook of Depression: Treatment, Assessment, and Research* (pp. 82-99). Homewood, IL: Dorsey Press.

Bhugra, D., & Mastrogianni, A. (2004). Globalization and mental disorders: Overview with relation to depression. *British Journal of Psychiatry, 184,* 10-20.

Blatt, S. J., & Zuroff, D. C. (2005). Empirical evaluation of the assumptions in identifying evidence based treatments in mental health. *Clinical Psychology Review, 25,* 459-486.

Bosc, M. (2000). Assessment of social functioning in depression. *Comprehensive Psychiatry, 41,* 63-69.

Bosc, M., Dubini, A., & Polin, V. (1997). Development and validation of a social functioning scale, the Social Adaptation Self-evaluation Scale. *European Neuropsychopharmacology, 7 (Supplement 1),* 57-70.

Bothwell, S. & Weissman, M. M. (1977). Social impairments four years after an acute depressive episode. *American Journal of Orthopsychiatry, 47*, 231-237.

Branch, L. G., & Jette, A. M. (1981). The Framingham Disability Study: I. Social disability among the aging. *American Journal of Public Health, 71,* 1202–1210.

Brayfield, A. H. (1965). Human effectiveness. *American Psychologist, 20,* 645-651.

Burlingame, G. M., Lambert, M. J., Reisinger, C. W., Neff, W. M., & Mosier, J. (1995). Pragmatics of tracking mental health outcomes in a managed care setting. *Journal of Mental Health Administration, 22,* 226-236.

Bushman, B. J., & Wells, G. L. (2001). Narrative impressions of literature: The availability bias and the corrective properties of meta-analytic approaches. *Personality and Social Psychology Bulletin, 27,* 1123-1130.

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

Campbell, D. T., & Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin, 56*, 81-105.

Cassano, P., & Fava, M. (2002). Depression and public health: an overview. *Journal of Psychosomatic Research, 53*, 849-857.

Cavanagh, K., Shapiro, D. A., Van Den Berg, S., Swain, S., Barkham, M., & Proudfoot, J. (2006). The effectiveness of computerized cognitive behavioural therapy in routine care. *British Journal of Clinical Psychology, 45,* 499-514.

Clark, D., Beck, A. T., & Alford, B. (1999). *Scientific Foundations of Cognitive Theory and Therapy of Depression.* New York, NY: John Wiley.

Clark, L. A., Vittengl, J. R., Dolores, K. & Jarrett, R. B. (2003). Shared, not unique, components of personality and psychosocial functioning predict depression severity after acute-phase cognitive therapy. *Journal of Personality Disorders, 17,* 406-430.

Clark, L. A., & Watson, D. (1991). Tripartite model of anxiety and depression: Psychometric evidence and taxonomic implications. *Journal of Abnormal Psychology, 100,* 316-336.

Clare, A. W., & Cairns, V. (1978). Design, development and use of a standardized interview to assess social maladjustment and dysfunction in community studies. *Psychological Medicine, 8*, 589-605.

Clare, A. W., Corney, R. H., & Cairns, V. E. (1984). Social adjustment: The design and use of an instrument for social work and social work research. *British Journal of Social Work, 14,* 323-336.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). New Jersey: Lawrence Erlbaum.

Cole, D. A., & Maxwell, S. E. (2003). Testing mediational models with longitudinal data: Questions and tips in the use of structural equation modeling. *Journal of Abnormal Psychology, 112,* 558-577.

Cortina, J. M. (1993). What is coefficient alpha? An examination of theory and application. *Journal of Applied Psychology, 78,* 98-104.

Coryell, W., Scheftner, W., Keller, M., Endicott, J., Maser, J., Klerman, G. L. (1993). The enduring psychosocial consequences of mania and depression. *American Journal of Psychiatry, 150,* 720-727.

Crowe, M. J. (1978). Conjoint marital therapy: A controlled outcome study. *Psychological Medicine, 8,* 623-636.

Crowe, M., & Luty, S. (2005). The process of change in Interpersonal Psychotherapy (IPT) for depression: A case study for the new IPT therapist. *Psychiatry: Interpersonal and Biological Processes, 68,* 43-54.

Cutrona, C. E., & Troutman, B. R. (1986). Social support, infant temperament, and parenting self-efficacy: A mediational model of postpartum depression. *Child Development, 57,* 1507-1518.

De Lisio, G., Maremmani, I., Perugi, G., Cassano, G. B., Deltito, J., & Akiskal, H. S. (1986). Impairment of work and leisure in depressed outpatients. A preliminary communication. *Journal of Affective Disorders, 10,* 79-84.

Della-Posta, C. & Drummond, P. D. (2006). Cognitive behavioural therapy increases re-employment of job seeking worker's compensation clients. *Journal of Occupational Rehabilitation, 16*, 223-230.

DeRubeis,R. J., Evans, M. D., Hollon, S. D., Garvey, M. J., Grove, W. M., Tuason, V. B., et al. (1990). How does cognitive therapy work? Cognitive change and symptom change in cognitive therapy and pharmacotherapy for depression. *Journal of Consulting and Clinical Psychology, 58*, 862-869.

Dickerson, F. B., Parente, F., & Ringel, N. (2000). The relationship among three measures of social functioning in outpatients with schizophrenia. *Journal of Clinical Psychology, 56*, 1509-1519.

Diener, E., Suh, E. M., Lucas, R. E., & Smith, H. L. (1999). Subjective well-being: Three decades of progress. *Psychological Bulletin, 125*, 276-302.

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

Dozois, D. J. A., Dobson, K. S., & Ahnberg, J. L. (1988). A psychometric evaluation of the Beck Depression Inventory-II. *Psychological Assessment, 10,* 83-89.

Dubini, A., Bosc, M., & Polin, V. (1997). Noradrenaline-selective versus serotonin-selective antidepressant therapy: Differential effects on social functioning. *Journal of Psychopharmacology, 11 (Supplement 4)*, 17-23.

Dunn, T. W., & Jarrett, R. B. (in press). Psychosocial Functioning in Depression. In R. Ingram (Ed.), *The International Encyclopedia of Depression*.

Dunn, T. W., Vittengl, J. & Jarrett, R. (2007, October). The effect of cognitive behavioral therapy on psychosocial functioning: A meta-analysis of adults with major depressive disorder. Poster presented at the Annual University of Texas Southwestern Medical Center Graduate Student Poster Session, Dallas, TX.

Dunner, D. L., Rush, A. J., Russell, J. M., Burke, M., Woodard, S., Wingard, P., et al. (2006). Prospective, long-term, multicenter study of the naturalistic outcomes of patients with treatment-resistant depression. *Journal of Clinical Psychiatry, 67,* 688–695.

Eisen, J. L., Mancebo, M. A., Pinto, A., Coles, M. E., Pagano, M. E., Stout, R., & Rasmussen, S. A. (2006). Impact of obsessive-compulsive disorder on quality of life. *Comprehensive Psychiatry, 47,* 270-275.

Elkin, I., Parloff, M. B., Hadley, S. W., & Autry, J. H. (1985). NIMH Treatment of Depression Collaborative Research Program: Background and research plan. *Archives of General Psychiatry, 42,* 305-316.

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

Emanuels-Zuurveen, L., & Emmelkamp, P. M. (1996). Individual behavioural-cognitive therapy v. marital therapy for depression in maritally distressed couples. *British Journal of Psychiatry, 169,* 181-188.

Emanuels-Zuurveen, L., & Emmelkamp, P. M. (1997). Spouse-aided therapy with depressed patients. *Behavior Modification, 21,* 62-77.

Endicott, J., Nee, J., Harrison, W., & Blumenthal, R. (1993). Quality of life enjoyment and satisfaction questionnaire: A new measure. *Psychopharmacology Bulletin, 29*, 321-326.

Endicott, J., Spitzer, R. L., Fleiss, J. L., & Cohen, J. (1976). The Global Assessment Scale: A procedure for measuring overall severity of psychiatric disturbance. *Archives of General Psychiatry, 33*, 766-771.

Epstein, N. B., Baldwin, L. M., & Bishop, D. S. (1983). The McMaster Family Assessment Device. *Journal of Marital & Family Therapy, 9,* 171-180.

Evans, M. D., Hollon, S. D., DeRubeis, R. J., Piasecki, J. M., Grove, W. M., Garvey, M. J., et al. (1992). Differential relapse following cognitive therapy and pharmacotherapy for depression. *Archives of General Psychiatry, 49,* 802-808.

Faravelli, C., Albanesi, G., & Poli, E. (1986). Assessment of depression: A comparison of rating scales. *Journal of Affective Disorders, 11*, 245-253.

Finkelstein, S. N., Berndt, E. R., Greenberg, P. E., Parsley, R. A., Russell, J. M., & Keller, M. B. (1996). Improvement in subjective work performance after

treatment of chronic depression: Some preliminary results. *Psychopharmacology Bulletin, 32*, 33–40.

First, M.B., Spitzer, R.L., Gibbon, M., & Williams, J.B. (1996). *Structured Clinical Interview for DSM-IV Axis I Disorders-Patient Edition* (SCID-I/P, Version 2.0). New York: New York State Psychiatric Institute, Biometrics Research Department.

Follette, W. C., & Greenberg, L. S. (2005). Technique factors in treating dysphoric disorders. In L. G. Castonguay & L. E. Beutler (Eds.), *Principles of Therapeutic Change that Work*, (pp. 83-110). New York: Oxford University Press.

Froyd, J. E., Lambert, M. J., & Froyd, J. D. (1996). A survey and critique of psychotherapy outcome measurement. *Journal of Mental Health, 5*, 11–15.

Garratt, G., Ingram, R. E., Rand, K. L., & Sawalani, G. (2007). Cognitive processes in cognitive therapy: Evaluation of the mechanisms of change in the treatment of depression. *Clinical Psychology: Science and Practice, 14,* 224-239.

Gelhart, R. P., & King, H. L. (2001). The influence of comorbid risk factors on the effectiveness of cognitive-behavioral treatment of depression. *Cognitive and Behavioral Practice, 8,* 18-28.

Giller, E., Bialos, D., Riddle, M. A., & Waldo, M. C. (1988). MAOI treatment response: multiaxial assessment. *Journal of Affective Disorders, 14,* 171-175.

Gloaguen, V., Cottraux, J., Cucherat, M., & Blackburn. I. (1998). A meta-analysis of the effects of cognitive therapy in depressed patients. *Journal of Affective Disorders, 49,* 59-72.

Greenberg, P. E., Kessler, R. C., Birnbaum, H. G., Leong, S. A., Lowe, S. W., Berglund, P. A., & Corey-Lisle, P. K. (2003). The economic burden of depression in the United States: How did it change between 1990 and 2000? *Journal of Clinical Psychiatry, 64,* 1465-1475.

Gold, P. W., & Chrousos, G. P. (2002). Organization of the stress system and its dysregulation in melancholic and atypical depression: High vs. low CRH/NE states. *Molecular Psychiatry, 7,* 254-275.

Goldman, H. H., Skodol, A. E., Lave, T. R. (1992). Revising axis V for DSM-IV: A review of measures of social functioning. *American Journal of Psychiatry, 149*, 1148-1156.

Gurland, B. J., Yorkston, N. J., Stone, A. R., Frank, J. D., & Fleiss, J. L. (1972). The Structured and Scaled Interview to Assess Maladjustment (SSIAM). I. Description, rationale, and development. *Archives of General Psychiatry, 27,* 259-264.

Haasen, C., & Sardashti, H. (2000). Relationship between depression and psychosocial stressors among Iranian emigrants. *Psychiatrische Praxis, 27,* 74–76.

Haby, M. M., Donnelly, M., Corry, J., & Vos, T. (2006). Cognitive behavioural therapy for depression, panic disorder and generalized anxiety disorder: a meta-regression of factors that may predict outcome. *Australian and New Zealand Journal of Psychiatry, 40,* 9-19.

Hamilton, M. (1960). A rating scale for depression. *Journal of Neurology, Neurosurgery and Psychiatry, 12,* 52-62.

Hammen, C. (2005). Stress and depression. *Annual Review of Clinical Psychology, 1,* 293-319.

Hasler, G., Moergeli, H., & Schnyder, U. (2004). Outcome of psychiatric treatment: What is relevant for our patients? *Comprehensive Psychiatry, 45,* 199-205.

Hays, R. D., Wells, K. B., Sherbourne, C. D., Rogers, W., & Spritzer, K. (1995). Functioning and well-being outcomes of patients with depression compared with chronic general medical illnesses. *Archives of General Psychiatry, 52,*11-19.

Healy, D., & McMonagle, T. (1997). The enhancement of social functioning as a therapeutic principle in the management of depression. *Journal of Psychopharmacology, 11 (Supplement 4),* 25-31.

Heim, C., & Nemeroff, C. B. (2001). The role of childhood trauma in the neurobiology of mood and anxiety disorders: Preclinical and clinical studies. *Biological Psychiatry, 49,* 1023-1039.

Henriksson, M. M., Aro, H. M., Marttunen, M. J., Heikkinen, M. E., Isometsa, E. T., Kuoppasalmi, K. I., et al. (1993). Mental disorders and comorbidity in suicide. *American Journal of Psychiatry, 150,* 935-940.

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

Hilsenroth, M. J., Ackerman, S. J., Blagys, M. D., Baumann, B. D., Baity, M. R., Smith, S. R., Price, J. L., Smith, C. L., Heindselman, T. L., Mount, M. K., Holdwick, D. J. (2000). Reliability and validity of DSM-IV axis V. *American Journal of Psychiatry, 157,* 1858-1863.

Hirschfeld, R.M.A., Dunner, D.L., Keitner, G., Klein, D.N., Koran, L.M., Kornstein, S.G., Markowitz, J.C., Miller, I., Nemeroff, C.B., Ninan, P.T., Rush, A.J., Schatzberg, A.F., Thase, M.E., Trivedi, M.H., Borian, F.E., Crits-Christoph, P., Keller, M.B. (2002). Does psychosocial functioning improve independent of depressive symptoms? A comparison of nefazodone, psychotherapy, and their combination. *Biological Psychiatry, 51,* 123-133.

Hirschfeld, R. M., Montgomery, S. A., Keller, M. B., Kasper, S., Schatzberg, A. F., Möller, H. J., Healy, D., Baldwin, D., Humble, M., Versiani, M., Montenegro, R., Bourgeois, M. (2000). Social functioning in depression: A review. *Journal of Clinical Psychiatry, 61*, 268-275.

Hirschfeld, R. M., Russell, J. M., Delgado, P. L., Fawcett, J., Friedman, R. A., Harrison, W. M., Koran, L. M., Miller, I. W., Thase, M. E., Howland, R. H., Connolly, M. A., & Miceli, R. J. (1998). Predictors of response to acute treatment of chronic and double depression with sertraline or imipramine. *The Journal of Clinical Psychiatry, 59,* 669-675.

Hoencamp, E., Haffmans, P. M. J., & Duivenvoorden, H. J. (1998). Predictability of dropout in unipolar depressed outpatients. *European Psychiatry, 13,* 63-66.

Hollon, S. D., & Beck, A. T. (2004). Cognitive and cognitive-behavioral therapies. In M. J. Lambert (Ed.), *Garfield and Bergin's handbook of psychotherapy and behavior change* (5th ed., pp. 447-492). New York: Wiley.

Hollon, S. D., DeRubeis, R. J., Evans, M. D., Wiemer, M. J., Garvey, M. J., Grove, W. M., & Tuason, V. B. (1992). Cognitive therapy and pharmacotherapy for

depression: Singly and in combination. *Archives of General Psychiatry, 49,* 774-781.

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

Hooley, J. M., & Teasdale, J. D. (1989). Predictors of relapse in unipolar depressives: Expressed emotion, marital distress, and perceived criticism. *Journal of Abnormal Psychology, 98,* 229-235.

Horowitz, L. M., Rosenberg, S. E., Baer, B. A., Ureño, G., & Villaseñor, V. S. (1988). Inventory of interpersonal problems: Psychometric properties and clinical applications. *Journal of Consulting and Clinical Psychology, 56,* 885-892.

*Imber, S. D., Pilkonis, P. A., Sotsky, S. M., Elkin, I., Watkins, J. T., Collins, J. F., Shea, M. T., Leber, W. R., & Glass, D. R. (1990). Mode-specific effects among three treatments for depression. *Journal of Consulting & Clinical Psychology, 58,* 352-359.

Ingram, R. E., Miranda, J., & Segal, Z. V. (1998). *Cognitive vulnerability to depression*. New York: Guilford Press.

*Jacobson, N. S., Dobson, K., Fruzzetti, A. E., Schmaling, K. B., Salusky, S. (1991). Marital therapy as a treatment for depression. *Journal of Consulting and Clinical Psychology, 59,* 547-557.

Judd, L. L., & Akiskal, H. S. (2000). Delineating the longitudinal structure of depressive illness: Beyond clinical subtypes and duration thresholds. *Pharmacopsychiatry, 33*, 3-7.

Judd, L. L., Akiskal, H. S., Schettler, P. J., Endicott, J., Leon, A. C., Solomon, D. A., et al. (2005). Psychosocial disability in the course of bipolar I and II disorders: A prospective, comparative, longitudinal study. *Archives of General Psychiatry, 62,* 1322-1330.

Judd, L. L., Akiskal, H. S., Zeller, P. J., Paulus, M., Leon, A. C., Maser, J. D., Endicott, J., et al. (2000). Psychosocial disability during the long-term course of unipolar major depressive disorder. *Archives of General Psychiatry, 57*, 375-380.

Kazdin, A. E. (1986). Comparative outcome studies of psychotherapy: Methodological issues and strategies. *Journal of Consulting and Clinical Psychology, 54*, 95-105.

Keller, M. B. (2003). Past, present, and future directions for defining optimal treatment outcome in depression: Remission and beyond. *JAMA, 289*, 3152-3160.

Keller, M. B., Lavori, P. W., Friedman, B., Nielsen, E., Endicott, J., McDonald-Scott, P., et al. (1987). The longitudinal interval follow-up evaluation: A comprehensive method for assessing outcome in prospective longitudinal studies. *Archives of General Psychiatry, 44,* 540-548.

Kennedy, S. H., Eisfeld, B. S. & Cooke, R. (2001). Quality of life: An important dimension in assessing the treatment of depression. *Journal of Psychiatry and Neuroscience, 26*, 23-28.

Kessler, R. C. (2003). Epidemiology of women and depression. *Journal of Affective Disorders, 74,* 5-13.

Kessler, R. C., Berglund, P., Demler, O., Jin, R., & Walters, E. E. (2005). Lifetime prevalence and age-of-onset distributions of *DSM-IV* disorders in the National Comorbidity Survey Replication. *Archives of General Psychiatry, 62,* 593-602.

Kessler, R. C., & Frank, R. G. (1997). The impact of psychiatric disorders on work loss days. *Psychological Medicine, 27,* 861-873.

Klerman, G. L. (1989). Depressive disorders: Further evidence for increased medical morbidity and impairment of social functioning. *Archives of General Psychiatry 46,* 856-858.

Kline, R. B. (2005). *Principles and Practice of Structural Equation Modeling.* New York: Guilford.

Knight, B. G., & Stre, D. (1999). Cognitive behavioral psychotherapy with older adults. *Clinical Psychology: Science and Practice, 6,* 188-203.

Koran, L. M., Hamilton, S. H., Hertzman, M., Meyers, B. S., Halaris, A. E., Tollefson, G. D., Downs, J. M., Folks, D. G., Jeste, D. V., Lazarus, L. W., et al. (1995). Predicting response to fluoxetine in geriatric patients with major depression. *Journal of Clinical Psychopharmacology, 15,* 421-427.

Kraemer, H.C., Stice, E., Kazdin, A., Offord, D., & Kupfer, D. (2001). How do risk factors work together? Mediators, moderators, independent, overlapping, and proxy risk factors. *American Journal of Psychiatry, 158*, 848-856.

Kupfer, D., & Frank, E. (1974). *The KDS-15: A marital questionnaire.* Pittsburgh, PA: Western Psychiatric Institute and Clinic, University of Pittsburgh.

Lambert, M. J., & Barley, D. E. (2002). Research summary on the therapeutic relationship and psychotherapy outcome. In J. C. Norcross (Ed.), *Psychotherapy*

relationships that work: Therapist contributions and responsiveness to patients. New York: Oxford University Press.

Lambert, M.J., & Ogles, B.M. (2004). The efficacy and effectiveness of psychotherapy. In M.J. Lambert (Ed.), *Bergin & Garfield's handbook of psychotherapy and behavior change* (5th ed., pp. 139–193). New York: Wiley.

Laroche, I., Hodgins, S., & Toupin, J. (1995). Correlation between symptoms and social adjustment in patients suffering from schizophrenia or major affective disorder. *Canadian Journal of Psychiatry, 40*, 27-34.

Latimer, P. R., & Sweet, A. A. (1984). Cognitive versus behavioral procedures in cognitive-behavior therapy: A critical review of the evidence. *Journal of Behavior Therapy & Experimental Psychiatry, 15,* 9-22.

Leichsenring, F., & Rabung, S. (2008). Effectiveness of long-term psychodynamic psychotherapy: A meta-analysis. *JAMA, 300,* 1551-1565.

Lenderking, W. R., Tennen, H., Nackley, J. F., Hale, M. S., Turner, R. R., Testa, M. A. (1999). The effects of venlafaxine on social activity level in depressed outpatients. *Journal of Clinical Psychiatry, 60,* 157-163.

Leon, A. C., Solomon, D. A., Mueller, T. I., Turvey, C. L., Endicott, J., & Keller, M. B. (1999). The Range of Impairment Functioning Tool (LIFE-RIFT): A brief measure of functional impairment. *Psychological Medicine, 29,* 869-878.

Lépine, J. P., Gastpar, M., Mendlewicz, J., & Tylee, A. (1997). Depression in the community: the first pan-European study DEPRES (Depression Research in European Society). *International Clinical Psychopharmacology, 12*, 19-29.

Lewinsohn, P. M. (1974). A behavioral approach to depression. In R. J. Friedman & M. M. Katz (Eds.), *The psychology of Depression: Contemporary Theory and Research* (pp. 157-185). New York: Wiley.

Lewinsohn, P. M., Hoberman, H. M., Teri, L., & Hautzinger, M. (1985). An integrative theory of depression. In S. Reiss & R. Bootzin (Eds.), *Theoretical Issues in Behavior Therapy* (pp. 331-359). New York: Academic.

Lipsey, M. W., & Wilson, D. B. (2001). *Practical meta-analysis.* Thousand Oaks, CA: Sage Publications.

Locke, H. J., & Wallace, K. M. (1959). Short marital adjustment and prediction tests: Their reliability and validity. *Marriage and Family Living, 21*, 251-255.

Luyten, P., Blatt, S. J., Houdenhove, B. V., & Corveleyn, J. (2006). Depression research and treatment: Are we skating to where the puck is going to be? *Clinical Psychology Review, 26,* 985-999.

Maher, M. J., Mora, P. A., & Leventhal, H. (2006). Depression as a predictor of perceived social support and demand: a componential approach using a prospective sample of older adults. *Emotion, 6,* 450-458.

Major, B., Cozzarelli, C., Sciacchitano, A. M., Cooper, M. L., Testa, M., & Mueller, P. M. (1990). Perceived social support, self-efficacy, and adjustment to abortion. *Journal of Personality and Social Psychology, 59,* 452-463.

Mallinckrodt, C. H., Watkin, J. G., Molenberghs, G., & Carroll, R. J. (2004). Choice of the primary analysis in longitudinal clinical trials. *Pharmaceutical Statistics, 3,* 161-169.

Marlowe, J. F. (2002). Depression's surprising toll on worker productivity. *Employee Benefits Journal, 27,* 16-21.

McDonald, R. P., & Ho, M. H. R. (2002). Principles and practice in reporting structural equation analyses. *Psychological Methods, 7,* 64-82.

McNair, D. M., & Lorr, M. (1964). Three kinds of psychotherapy goals. *Journal of Clinical Psychology, 20,* 390-393.

Mee, J., & Sumsion, T. (2001). Mental health clients confirm the motivating power of occupation. *The British Journal of Occupational Therapy, 64*, 121-128.

Meyer-Lindenberg, A., Buckholtz, J. W., Kolachana, B. R., Hariri, A., Pezawas, L., Blasi, G., Wabnitz, A., Honea, R., Verchinski, B., Callicott, J. H., Egan, M., Mattay, V., & Weinberger, D. R. (2006). Neural mechanisms of genetic risk for impulsivity and violence in humans. *Proceedings of the National Academy of Sciences, 103,* 6269-6274.

Miklowitz, D. J., Otto, M. W., Frank, E., Reilly-Harrington, N. A., Wisniewski, S. R., Kogan, J. N., et al. (2007). Psychosocial treatments for bipolar depression: A 1-year randomized trial from the Systematic Treatment Enhancement Program.*Archives of General Psychiatry, 64,* 419-427.

Miller, I. W., Keitner, G. I., Schatzberg, A. F., Klein, D. N., Thase, M. E., Rush, A. J., Markowitz, J. C., Schlager, D. S., Kornstein, S. G., Davis, S. M., Harrison, W. M., & Keller, M. B. (1998). The treatment of chronic depression, part 3: Psychosocial functioning before and after treatment with sertraline or imipramine. *Journal of Clinical Psychiatry, 59*, 608-619.

Miller, N. E., & Dollard, J. (1941). *Social Learning and Imitation*. New Haven, Connecticut: Yale University Press.

Minino, A. M., Heron, M. P., Smith, B. L. (2006). Preliminary data for 2004. *National vital statistics reports, 54(19)*. Hyattsville, MD: National Center for Health Statistics.

Minium, E. W., King, B. W., & Bear, G. (1993). *Statistical reasoning in psychology and education* (3rd ed.). New York: Wiley.

Mintz, J., Mintz, L. I., Arruda, M. J., & Hwang, S. S. (1992). Treatments of depression and the functional capacity to work. *Archives of General Psychiatry, 49,* 761-768.

Miranda, J., Chung, J. Y., Green, B. L., Krupnick, J., Siddique, J., Revicki, D. A., & Belin, T. (2003). Treating depression in predominantly low-income young minority women: A randomized controlled trial. *JAMA, 290,* 57-65.

Mischel, W. (1973). Toward a cognitive social learning reconceptualization of personality. *Psychological Review, 80,* 252-283.

Montgomery, S. A., & Asberg, M. (1979). A new depression scale designed to be sensitive to change. *British Journal of Psychiatry, 134*, 382-389.

Moos, R. H., & Cronkite, R. C. (1999). Symptom-Based Predictors of a 10-Year Chronic Course of Treated Depression. *Journal of Nervous & Mental Disease, 187,* 360-368.

Mueller, T. I., Leon, A. C., Keller, M. B., Solomon, D. A., Endicott, J., Coryell, W., Warshaw, M., & Maser, J. D. (1999). Recurrence after recovery from major

depressive disorder during 15 years of observational follow-up. *American Journal of Psychiatry. 156,* 1000-1006.

Mulder, R. T., Joyce, P. R., & Frampton, C. (2003). Relationships among measures of treatment outcome in depressed patients. *Journal of Affective Disorders, 76,* 127-135.

Mulder, R. T., Joyce, P. R., Frampton, C. M., Luty, S. E., & Sullivan, P. F. (2006). Six months of treatment for depression: Outcome and predictors of the course of illness. *American Journal of Psychiatry, 163,* 95-100.

Mundt, J. C., Marks, I. M., Shear, M. K., Greist, J. H. (2002). The Work and Social Adjustment Scale: A simple measure of impairment in functioning. *British Journal of Psychiatry, 180,* 461-464.

*Murray, L., Cooper, P. J., Wilson, A., & Romaniuk, H. (2003) Controlled trial of the short-and long-term effect of psychological treatment of postpartum depression 2: Impact on the mother-child relationship and child outcome. *British Journal of Psychiatry, 182*, 420-427.

National Institute of Mental Health. (2000). *Translating Behavioral Science into Action: Report of the National Advisory Mental Health Council Behavioral Science Workgroup* (NIH Publication 00-4699). Rockville, MD: Author.

Nofzinger, E. A., Thase, M. E., Reynolds, C. F., Frank, E., Jennings, J. R., Garamoni, G. L., Fasiczka, A. L., & Kupfer, D. J. (1993). Sexual function in depressed men: Assessment by self-report, behavioral, and nocturnal penile tumescence measures before and after treatment with cognitive behavior therapy. *Archives of General Psychiatry, 50,* 24-30.

Oatley, K., & Bolton, W. (1985). A social cognitive theory of depression in reaction to life events. *Psychological Review, 92,* 372-388.

O'Connor, R. (2003). An integrative approach to treatment of depression. *Journal of Psychotherapy Integration, 13,* 130-170.

Ogrodniczuk, J. S., Piper, W. E., & Joyce, A. S. (2004). Residual symptoms in depressed patients who successfully respond to short-term psychotherapy. *Journal of Affective Disorders, 82,* 469-473.

Ormel, J., Oldehinkel, A. J., Nolen, W. A., Vollebergh, W. (2004). Psychosocial disability before, during, and after a major depressive episode: a 3-wave population-based study of state, scar, and trait effects. *Archives of General Psychiatry, 61,* 387-392.

Ormel, J., Von Korff, M., Van den Brink, W., Katon, W., Brilman, E., & Oldehinkel, T. (1993). Depression, anxiety, and social disability show synchrony of change in primary care patients. *American Journal of Public Health, 83,* 385-390.

Oquendo, M. A., Ellis, S. P., Greenwald, S., Malone, K. M., Weissman, M. M., Mann, J. J., et al. (2001). Ethnic and sex differences in suicide rates relative to major depression in the United States. *American Journal of Psychiatry, 158,* 1652-1658.

Papakostas, G. I., Petersen, T., Denninger, J. W., Tossani, E., Pava, J. A., Alpert, J. E., Nierenberg, A. A., & Fava, M. (2004). Psychosocial functioning during the treatment of major depressive disorder with fluoxetine. *Journal of Clinical Psychopharmacology, 24,* 507-511.

Papakostas, G. I., Petersen, T., Mahal, Y., Mischoulon, D., Nierenberg, A. A., & Fava, M. (2004). Quality of life assessments in major depressive disorder: A review of the literature. *General Hospital Psychiatry, 26,* 13-17.

Parsons, T. (1958). Definitions of health and illness in the light of American values and social structure. In E. G. Jaco (Ed.), *Patients, Physicians, and Illness* (pp. 120-144). New York: Free Press.

Paykel, E. S. (1999). Social functioning and the depressed patient. *International Journal of Psychiatry in Clinical Practice, 3,* S9-S11.

Paykel, E. S., & Weissman, M. M. (1973). Social adjustment and depression: A longitudinal study. *Archives of General Psychiatry, 28,* 659-663.

Paykel, E. S., Weissman, M. M. & Prusoff, B. A. (1978). Social maladjustment and severity of depression. *Comprehensive Psychiatry, 19,* 121-128.

Pedersen, R. D., Pallay, A. G., & Rudolph, R. L. (2002). Can improvement in well-being and functioning be distinguished from depression improvement in antidepressant clinical trials? *Quality of Life Research, 11*, 9-17.

Pollock, B. G., Perel, J. M., Kupfer, D. J., Bowler, K. A., & Miewald, J. M. (1993). Early response patterns associated with successful clomipramine treatment. *Journal of Clinical Psychopharmacology, 13,* 442-447.

Prince, M. J., Harwood, R. H., Thomas, A., & Mann, A. H. (1998). A prospective population-based cohort study of the effects of disablement and social milieu on the onset and maintenance of late-life depression. The Gospel Oak Project VII. *Psychological Medicine, 28,* 337-350.

*Propst, L. R., Ostrom, R., Watkins, P., Dean, T., & Mashburn, D. (1992). Comparative efficacy of religious and nonreligious cognitive-behavioral therapy for the treatment of clinical depression in religious individuals. *Journal of Consulting and Clinical Psychology, 60,* 94-103.

Proudfoot, J., Ryden, C., Everitt, B., Shapiro, D. A., Goldberg, D., Mann, A., Tylee, A., Marks, I., & Gray, J. A. (2004). Clinical efficacy of computerised cognitive-behavioural therapy for anxiety and depression in primary care: Randomised controlled trial. *British Journal of Psychiatry, 185,* 46-54.

*Rehm, L. P., Kaslow, N. J., & Rabin, A. S. (1987). Cognitive and behavioral targets in a self-control therapy program for depression. *Journal of Consulting and Clinical Psychology, 55,* 60-67.

Revicki, D. A., Turner, R., Brown, R., & Martindale, J. J. (1992). Reliability and validity of a health-related quality of life battery for evaluating outpatient antidepressant treatment. *Quality of Life Research, 1,* 257-266.

Reynolds, C. F., Frank, E., Thase, M. E., Houck, P. R., Jennings, J. R., Howell, J. R., Lilienfeld, S. O., & Kupfer, D. J. (1988). Assessment of sexual function in depressed, impotent, and healthy men: Factor analysis of a brief sexual function questionnaire for men. *Psychiatry Research, 24*, 231–225.

Riskind, J. H., Beck, A. T., Berchick, R. J., Brown, G., & Steer, R. A. (1987). Reliability of DSM-III diagnoses for major depression and generalized anxiety disorder using the Structured Clinical Interview for DSM-III. *Archives of General Psychiatry, 44,* 817-820.

Riso, L. P., Thase, M. E., Howland, R. H., Friedman, E. S., Simons, A. D., Tu, X. M. (1997). A prospective test of criteria for response, remission, relapse, recovery, and recurrence in depressed patients treated with cognitive behavior therapy. *Journal of Affective Disorders, 43,* 131-142.

Ro, E., Clark, L. A., Vittengl, J., & Jarrett, R. B., (2007, October). *Social Adjustment Scale in Patients with Recurrent Depression.* Poster presented at the 21st Annual Meeting of the Society for Research in Psychopathology, Iowa City, IA.

Rodriguez, B. F., Bruce, S. E., Pagano, M. E., & Keller, M. B. (2005). Relationships among psychosocial functioning, diagnostic comorbidity, and the recurrence of generalized anxiety disorder, panic disorder, and major depression. *Journal of Anxiety Disorders, 19*, 752-766.

Roth, A., & Fonagy, P. (2004) *What works for whom? A critical review of psychotherapy research* (2nd ed.). New York: Guilford Press.

Rounsaville, B. J., Prusoff, B. A., & Weissman, M. M. (1980). The course of marital disputes in depressed women: a 48-month follow-up study. *Comprehensive Psychiatry, 21*, 111-118.

Rounsaville, B. J., Weissman, M. M, & Prusoff, B. A. (1981). Psychotherapy with depressed outpatients: Patient and process variables as predictors of outcome. *The British Journal of Psychiatry, 138,* 67-74.

Rothschild, A. J., Samson, J. A., Bond, T. C., Luciana, M. M., Schildkraut, J. J., & Schatzberg, A. F. (1993). Hypothalamic-pituitary-adrenal axis activity and 1-year outcome in depression. *Biological Psychiatry, 34,* 392-400.

Rowe, H. (1997). *Work Potential Profile.* Melbourne: ACER.

Rush, A.J., Giles, D.E., Schlesser, M.A., Fulton, C.L., Weissenburger, J.E., & Burns, C.T. (1986). The Inventory for Depressive Symptomatology (IDS): Preliminary findings. *Psychiatry Research, 18,* 65-87.

Rush, A. J., Gullion, C. M., Basco, M. R., Jarrett, R. B., & Trivedi, M. H. (1996). The Inventory of Depressive Symptomatology (IDS): Psychometric properties. *Psychological Medicine, 26,* 477-486.

Rush, A. J., Kraemer, H. C., Sackeim, H. A., Fava, M., Trivedi, M. H., Frank, E., et al. (2006). Report by the ACNP task force on response and remission in major depressive disorder. *Neuropsychopharmacology, 31,* 1841-1853.

Saeki, T., Asukai, N., Miyake, Y., Miguchi, M., & Yamawaki, S. (2002). Characteristics of family functioning in patients with endogenous monopolar depression. *Hiroshima Journal of Medical Sciences, 51,* 55-62.

*Saulsman, L. M., Coall, D. A., & Nathan, P. R. (2006). The association between depressive personality and treatment outcome for depression following a group cognitive-behavioral intervention. *Journal of Clinical Psychology, 62*, 1181-1196.

Schmitt, N. (1996). Uses and abuses of coefficient alpha. *Psychological Assessment, 8,* 350-353.

Schotte, C. K. W., Bossche, B. V. D., Doncker, D. D., Claes, S., & Cosyns, P. (2006). A biopsychosocial model as a guide for psychoeducation and treatment of depression. *Depression and Anxiety, 23,* 312-324.

*Shapiro, D. A., Barkham, M., Rees, A., Hardy, G. E, Reynolds, S., & Startup, M. (1994). Effects of treatment duration and severity of depression on the effectiveness of cognitive-behavioral and psychodynamic-interpersonal psychotherapy. *Journal of Consulting and Clinical Psychology, 62,* 522-534.

Simons, A. D., Levine, J. L., Lustman, P. J., & Murphy, G. E. (1984). Patient attrition in a comparative outcome study of depression: A follow-up report. *Journal of Affective Disorders, 6*, 163-173.

Simon, G. E. (2003). Social and economic burden of mood disorders. *Biological Psychiatry, 54,* 208-215.

Skodol, A. E., Link, B. G., Shrout, P. E., & Horwath, E. (1988). The revision of axis V in DSM-III-R: Should symptoms have been included? *American Journal of Psychiatry, 145,* 825-829.

Sobel, M. E. (1986). Some new results on indirect effects and their standard errors in covariance structure models. In N. Tuma (Ed.), *Sociological Methodology 1986* (pp. 159-186). Washington, DC: American Sociological Association.

Solomon, D. A., Keller, M. B., Leon, A. C., Mueller, T. I., Lavori, P. W., Shea, M. T., Coryell, W., Warshaw, M., Turvey, C., Maser, J. D., & Endicott, J. (2000). Multiple recurrences of major depressive disorder. *American Journal of Psychiatry, 157*, 229-233.

Solomon, D. A., Leon, A. C., Endicott, J., Mueller, T. I., Coryell, W., Shea, M. T., & Keller, M. B. (2004). Psychosocial impairment and recurrence of major depression. *Comprehensive Psychiatry, 45,* 423-430.

Sotsky, S. M., Glass, D. R., Shea, M. T., Pilkonis, P. A., Collins, J. F., Elkin, I., Watkins, J. T., Imber, S. D., Leber, W. R., & Moyer, J. (1991). Patient predictors of response to psychotherapy and pharmacotherapy: findings in the NIMH Treatment of Depression Collaborative Research Program. *The American Journal of* Psychiatry, 148, 997-1008.

Spanier, G. B. (1976). Measuring dyadic adjustment: New scales for assessing the quality of marriage and similar dyads. *Journal of Marriage and the Family, 38,* 15-28.

Spijker, J., de Graaf, R., Bijl, R. V., Beekman, A. T. F., Ormel, J., & Nolen, W. A. (2004) Functional disability and depression in the general population. Results from the Netherlands Mental Health Survey and Incidence Study (NEMESIS). *Acta Psychiatrica Scandinavica 110*, 208-214.

Spitzer, R. L., Endicott, J., & Robins, E. (1975). *Research diagnostic criteria.* New York: Biometrics Research.

Spitzer, R. L., & Forman, J. B. (1979). DSM-III field trials: II. Initial experience with the multiaxial system. *American Journal of Psychiatry, 136,* 818-820.

Spitzer, R. L., & Wakefield, J. C. (1999). DSM-IV diagnostic criterion for clinical significance: does it help solve the false positives problem? *American Journal of Psychiatry, 156,* 1856-1864.

Stewart, W. F., Ricci, J. A., Chee, E., Hahn, S. R., & Morganstein, D. (2003). Cost of lost productive work time among US workers with depression. *JAMA, 289,* 3135-3144.

Sue, S., Fujino, D.C., Hu, L., Takeuchi, D. T., & Zane, N. W.S. (1991). Community mental health services for ethnic minority groups: A test of the cultural

responsiveness hypothesis. *Journal of Consulting and Clinical Psychology, 59,* 533-540.

Sullivan, M. J., Adams, H., Thibault, P., Corbière, M., & Stanish, W. D. (2006). Initial depression severity and the trajectory of recovery following cognitive-behavioral intervention for work disability. *Journal of Occupational Rehabilitation, 16,* 63-74.

Sussman, L. K., Robins, L. N., & Earls, R. (1987). Treatment seeking for depression by black and white Americans. *Social Science Medicine, 24,* 187-196.

Tanner, J., Weissman, M., & Prusoff, B. (1975). Social adjustment and clinical relapse in depressed outpatients. *Comprehensive Psychiatry, 16,* 547-556.

Thase, M. E. (2003). Evaluating antidepressant therapies: Remission as the optimal outcome. *Journal of Clinical Psychiatry, 64 (Supplement 13),* 18-25.

*Thase, M. E., Bowler, K., & Harden, T. (1991). Cognitive behavior therapy of endogenous depression: Part 2. Preliminary findings in 16 unmedicated inpatients. *Behavior Therapy, 22,* 469–477.

*Thase, M. E., Reynolds, C. F., Frank, E., Jennings, J. R., Nofzinger, E., Fasiczka, A. L., Garamoni, G., & Kupfer, D. J. (1994). Polysomnographic studies of unmedicated depressed men before and after cognitive behavioral therapy. *American Journal of Psychiatry, 151,* 1615-1622.

Thase, M. E., Simons, A. D., Cahalane, J. F., & McGeary, J. (1991). Cognitive behavior therapy of endogenous depression: I. An outpatient clinical replication series. *Behavior Therapy, 22,* 457-467.

*Thase, M. E., Simons, A. D., Cahalane, J. F., McGeary, J., Harden, T. (1991). Severity of depression and response to cognitive behavior therapy. *American Journal of Psychiatry, 148,* 784-789.

*Thompson, L. W., Gallagher, D., & Breckenridge, J. S. (1987). Comparative effectiveness of psychotherapies for depressed elders. *Journal of Consulting and Clinical Psychology, 55,* 385-390.

Trivedi, M. H. (2001). Sensitizing clinicians and patients to the social and functional aspects of remission. *Journal of Clinical Psychiatry. 62 (Supplement 19),* 32-35.

Trivedi, M. H., Morris, D. W., Grannemann, B. D., & Mahadi, S. (2005). Symptom clusters as predictors of late response to antidepressant treatment. *Journal of Clinical Psychiatry, 66,* 1064-1070.

Trivedi, M. H., Rush, A. J., Ibrahim, H. M., Carmody, T. J., Biggs, M. M., Suppes, T., Crismon, M. L., Shores-Wilson, K., Toprac, M. G., Dennehy, E. B., Witte, B., & Kashner, T. M. (2004). The Inventory of Depressive Symptomatology, Clinician Rating (IDS-C) and Self-Report (IDS-SR), and the Quick Inventory of Depressive Symptomatology, Clinician Rating (QIDS-C) and Self-Report (QIDS-SR) in public sector patients with mood disorders: A psychometric evaluation. *Psychological Medicine, 34,* 73-82.

Trivedi, M. H., Rush, A. J., Wisniewski, S. R., Warden, D., McKinney, W., Downing, M., Berman, S. R., Farabaugh, A., Luther, J. F., Nierenberg, A. A., Callan, J. A., & Sackeim, H. A. (2006). Factors associated with health-related quality of life among outpatients with major depressive disorder: A STAR*D report. *Journal of Clinical Psychiatry, 67,*185-195.

Tylee, A., Gastpar, M., Lépine, J. P., & Mendlewicz, J. (1999). DEPRES II (Depression Research in European Society II): A patient survey of the symptoms, disability and current management of depression in the community. DEPRES Steering Committee. *International Clinical Psychopharmacology, 14,* 139-151.

Tyrer, P. J. (1990). Personality disorder and social functioning. In D. F. Peck & C. M. Shapiro (Eds.), *Measuring human problems: A practical guide* (pp. 119-142). Chichester, UK: Wiley.

Tyrer, P. J., & Casey, P. R. (Eds.). (1993). *Social function in psychiatry: The hidden axis of classification exposed*. Bristol, PA: Wrightson Biomedical Publication.

U.S. Census Bureau. (2002). *Census 2000 Modified Race Data* [MR(31)-CO.txt].

*Vittengl, J. R., Clark, L. A., & Jarrett, R. B. (2004). Improvement in social-interpersonal functioning after cognitive therapy for recurrent depression. *Psychological Medicine, 34,* 643-658.

Vittengl, J. R., Clark, L. A., & Jarrett, R. B. (2008). Deterioration in psychosocial functioning predicts relapse/recurrence after cognitive therapy for depression. *Journal of Affective Disorder.*

Vittengl, J. R., Clark, L. A., Kraft, D., & Jarrett, R. B. (2005). Multiple measures, methods, and moments: A factor-analytic investigation of change in depressive symptoms during acute-phase cognitive therapy for depression. *Psychological Medicine, 35,* 693-704.

Von Korff, M., Ormel, J., Katon, W., & Lin, E. H. (1992). Disability and depression among high utilizers of health care. A longitudinal analysis. *Archives of General Psychiatry, 49*, 91-100.

Wampold, B. E., Mondin, G. W., Moody, M., Stich, F., Benson, K., & Ahn, H. (1997). A meta-analysis of outcome studies comparing bona fide psychotherapies: Empirically, "All must have prizes". *Psychological Bulletin, 122,* 203-215.

Ward, E., King, M., Lloyd, M., Bower, P., Sibbald, B., Farrelly, S., Gabbay, M., Tarrier, N., & Addington-Hall, J. (2000). Randomised controlled trial of non-directive counselling, cognitive-behaviour therapy, and usual general practitioner care for patients with depression. I: Clinical effectiveness. *BMJ, 321*, 1383-1388.

Ware, J. E., & Sherbourne, C. D. (1992). The MOS 36-item short-form health survey (SF-36). I. Conceptual framework and item selection. *Medical Care, 30*, 473-483.

Warshaw, M. G., Dyck, I., Allsworth, J., Stout, R. L., & Keller, M. B. (2001). Maintaining reliability in a long-term psychiatric study: An ongoing inter-rater reliability monitoring program using the longitudinal interval follow-up evaluation. *Journal of Psychiatric Research, 35,* 297-305.

Warshaw, M. G., Keller, M. B., & Strout, R. L. (1994). Reliability and validity of the longitudinal interval follow-up evaluation for assessing outcome of anxiety disorders. *Journal of Psychiatric Research, 28,* 531-545.

Weissman, M. M. (2000). Social functioning and the treatment of depression. *Journal of Clinical Psychiatry, 61(Supplement 1),* 33-38.

Weissman, M. M., Bland, R. C., Canino, G. J., Faravelli, C., Greenwald, S., Hwu, H. G., Weissman, M. M., & Bothwell, S. (1976). Assessment of social adjustment by patient self-report. *Archives of General Psychiatry, 33,* 1111-1115.

Weissman, M. M., Klerman, G. L., Prusoff, B. A., Hanson, B., & Paykel, E. S. (1976). The efficacy of psychotherapy in depression: symptom remission and response to treatment. *Proceedings of the Annual Meeting of the American Psychopathological Association, 64,* 165-177.

Weissman, M. M., Markowitz, J. C., & Klerman, G. L. (2000). *Comprehensive Guide to Interpersonal Psychotherapy*. New York: Basic Books.

Weissman, M. M., Olfson, M., Gameroff, M. J., Feder, A., & Fuentes, M. (2001). A comparison of three scales for assessing social functioning in primary care. *American Journal of Psychiatry, 158,* 460-466.

Weissman, M. M., Paykel, E. S., Siegel, R., & Klerman, G. L. (1971). The social role performance of depressed women: comparisons with a normal group. *American Journal of Orthopsychiatry, 41,* 390-405.

Weissman, M. M., Prusoff, B. A., Thompson, D., Harding, P. S., & Myers, J. K. (1978). Social adjustment by self-report in a community sample and in psychiatric outpatients. *Journal of Nervous and Mental Disease, 166,* 317-326.

Weissman, M. M., Sholomskas, D., & John, K. (1981). The assessment of social adjustment: An update. *Archives of General Psychiatry, 38,* 1250-1258.

Whisman, M. A. (1993). Mediators and moderators of change in cognitive therapy of depression. *Psychological Bulletin, 114,* 248-265.

Wiersma, D. (1996). Measuring social disabilities in mental health. *Social Psychiatry and Psychiatric Epidemiology, 31,* 101-108.

Wiersma, D., DeJong, A., & Ormel, J. (1988). The Groningen Social Disabilities Schedule: development, relationship with I.C.I.D.H., and psychometric properties. *International Journal of Rehabilitation Research, 11*, 213-224.

Williams, R. A., & Strasser, P. B. (1999). Depression in the workplace: Impact on employees. *American Association of Occupational Health Nurses Journal, 47,* 526-537.

Wilson, D. B. (2001). *Effect size determination program.* College Park: University of Maryland.

Wilson, G. T., Fairburn, C. C., Agras, W. S., Walsh, B. T., & Kraemer, H. (2002). Cognitive-behavioral therapy for bulimia nervosa: time course and mechanisms of change. *Journal of Consulting and Clinical Psychology, 70,* 267-274.

Winokur, G. (1979). Familial (genetic) subtypes of pure depressive disease. *American Journal of Psychiatry, 136,* 911-913.

World Health Organization. (1980). *International Classification of Impairments, Disabilities, and Social Handicaps (ICIDH).* Geneva, Switzerland: World Health Organization.

Young, J., Beck, A. T. (1980). *Cognitive Therapy Scale: Rating Manual.* Philadelphia, Pa: Center for Cognitive Therapy.

Zimmerman, M., McGlinchey, J., Posternak, M., Friedman, M., Boerescu, D., & Attiullah, N. (2006). Discordance between self-reported symptom severity and psychosocial functioning ratings in depressed outpatients: Implications for how

remission from depression should be defined. *Psychiatry Research, 141*, 185-191.